

Systematic Review

Behavioral Biometric Continuous Authentication for Mobile Devices with an Intelligent Personal Agent: A Systematic Review

Madi Gali ^{1,*}, Aray Kassenkhan ², Yersain Chinibayev ², Aigerim Abshukirova ³ and Vassiliy Serbin ^{4,*}

¹ Department of Cybersecurity, Institute of Automation and Information Technologies, Satbayev University, Almaty 050013, Kazakhstan

² Department of Software Engineering, Institute of Automation and Information Technologies, Satbayev University, Almaty 050013, Kazakhstan; a.kassenkhan@satbayev.university (A.K.); y.chinibayev@satbayev.university (Y.C.)

³ National Scientific Laboratory of Collective Use of Information and Space Technologies, Institute of Automation and Information Technologies, Satbayev University, Almaty 050013, Kazakhstan; a.abshukirova@satbayev.university

⁴ Department of Information Systems, Institute of Automation and Information Technologies, Satbayev University, Almaty 050013, Kazakhstan

* Correspondence: gali.m@stud.satbayev.university (M.G.); v.serbin@satbayev.university (V.S.)

Abstract

Static, one-time authentication mechanisms such as passwords and PINs are increasingly inadequate for protecting mobile devices throughout an active session. Behavioral biometric continuous authentication (BBCA) addresses this gap by passively monitoring user-specific interaction patterns—keystroke dynamics, touch and swipe gestures, gait, and motion—to verify identity on an ongoing basis. This systematic review synthesizes 80 studies selected via a PRISMA-compliant protocol from IEEE Xplore, ACM Digital Library, Scopus, ScienceDirect, Web of Science, and SpringerLink (2017–2025). We examine behavioral and multimodal biometric modalities, machine learning approaches ranging from classical classifiers to deep sequence and transformer architectures, and their integration with intelligent personal agents, wearable devices, and IoT/edge infrastructures. Security analyses cover spoofing, adversarial and generative attacks, mimicry, and model-level threats including membership inference and reconstruction. Privacy-preserving mechanisms—cancelable biometrics, Bloom filter encodings, zero-knowledge proof protocols, federated learning, and blockchain-based identity management—are evaluated against practical trade-offs in energy consumption and latency on resource-constrained devices. Key research gaps are identified: the absence of standardized adversarial benchmarks, lack of end-to-end pipeline evaluations under simultaneous adversarial and privacy threat models, and limited user-centered studies on consent and acceptance of privacy-preserving mechanisms under frameworks such as GDPR. Recommended future directions combine adaptive multimodal fusion, privacy-preserving cryptography, energy-aware modality selection, and interdisciplinary human-centered evaluation to advance practical, resilient continuous authentication for mobile and assistant-enriched environments.

Keywords: continuous authentication; behavioral biometrics; keystroke dynamics; gait recognition; touch dynamics; intelligent personal agent; federated learning; privacy-preserving authentication; mobile security; IoT; zero-knowledge proof; adversarial robustness



Academic Editor: Loris Nanni

Received: 9 June 2026

Revised: 8 July 2026

Accepted: 15 July 2026

Published: 22 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

1. Introduction

Continuous authentication has emerged in response to the shortcomings of static login mechanisms that verify a user only once at the beginning of a session. When a device or service remains authenticated after the initial check, any subsequent loss or theft can expose sensitive data, since passwords, PINs, or one-time checks alone are insufficient to safeguard modern smartphones and connected systems. Behavioral biometric continuous authentication (BBCA) extends protection beyond the login point by monitoring user behavior throughout a session and repeatedly validating identity, thereby aiming to provide a safer alternative to traditional mechanisms in mobile contexts. Behavioral biometrics exploit observable actions and interaction patterns, such as walking gait, touch gestures, keystroke dynamics, and device usage profiles, rather than relying on static physiological traits alone [1]. On smartphones, built-in sensors and interaction logs enable the capture of motion behavior, touch dynamics, and keystroke patterns to distinguish legitimate users from impostors during everyday activities like reading or app navigation [2,3]. Studies have also explored combinations of touch events with accelerometer, gyroscope, and magnetometer readings to implement continuous user authentication during passive tasks, reporting highly effective behavioral biometric performance in smartphone scenarios [2]. These developments align with broader efforts to deploy sensor-based behavioral biometrics as an indirect, continuous validation mechanism on mobile devices [3].

Beyond pure behavioral signals, continuous authentication systems frequently incorporate physiological biometrics, including fingerprints and iris patterns, alongside behavioral modalities such as keystroke dynamics and gait. Physiological traits typically provide strong one-time authentication, whereas behavioral signals support passive and ongoing checks that can improve usability during prolonged sessions. Reviews of continuous authentication systems (CAS) emphasize that combining physiological, behavioral, and multimodal approaches can yield robust and versatile authentication frameworks that better balance accuracy, continuity, and user comfort [4]. Multimodal fusion strategies have been repeatedly associated with improved results compared to single modality schemes, and are widely regarded as a future trend for securing mobile and smartphone environments [1,4].

At the same time, research on intelligent personal assistants (IPAs) exposes complementary security and privacy concerns that motivate stronger authentication safeguards on mobile and ambient platforms. Commercial and academic IPA solutions may be vulnerable to remote manipulation of voice activation, making it possible to control devices through spoofed or unauthorized voice commands. Such assistants often depend on the persistent storage of sensitive information, including real-time location and speech data, which raises substantial privacy risks for end users. Systematic analyses of IPA applications further identify a lack of standardized measures for comparing assistant performance and security, and report broad, unexplored opportunities in domains such as education, health, and privacy where personal and cognitive assistants could support ambient assisted living and smart environments [5].

Within the broader security landscape, mobile networks and the Internet of Things (IoT) intensify the relevance of continuous and context-aware authentication. Early cellular generations suffered from fundamental security flaws, including weak or absent encryption, one-way authentication, and susceptibility to unauthorized base stations, which exposed users to eavesdropping and data theft [6]. Contemporary IoT deployments face challenges of mobility, scalability, and pervasive cyber-physical threats, making device and service level authentication essential for maintaining confidentiality, integrity, and availability. Trust based monitoring schemes with intelligent agents and middleware have been proposed to strengthen IoT communication security, reduce misdetection and false

positives, and improve network lifetime via energy efficiency, underscoring the need for lightweight yet reliable security primitives in resource-constrained environments [7]. In healthcare-oriented IoT settings, lightweight authentication and key agreement protocols have been designed to lower computational and communication overhead while enhancing security, with layered, federated learning-based mechanisms further aiming to preserve device privacy and maintain high authentication accuracy on standard datasets [8].

Continuous authentication research must also contend with significant usability, security, privacy, and ethical issues. Analyses of Continuous User Biometric Authentication Systems (CUBAS) highlight a persistent trade-off between usability and security, as excessively strict thresholds increase False Rejections and user frustration, whereas relaxed thresholds risk unauthorized access. Many behavioral and context-aware modalities remain susceptible to spoofing and adversarial attacks, with generative techniques capable of elevating error rates in touch and keystroke-based systems. Persistent monitoring through sensors consumes energy on mobile devices and may strain battery life, especially when multiple sensors operate continuously. Privacy implications arise from the continuous collection and storage of biometric and contextual data, often without transparent user awareness or explicit consent regarding data usage and retention. Ethical considerations further demand careful treatment of consent, potential bias, and user autonomy when deploying behavioral biometric systems in real-world environments [4].

Existing surveys provide important foundations but also reveal specific gaps that justify a focused review at the intersection of continuous mobile authentication, behavioral biometrics, and intelligent personal agents. Systematic reviews on sensor-based continuous authentication for smartphones categorize over 140 works into motion, gait, keystroke dynamics, gesture, voice, and multimodal groups, outlining system designs, enrollment procedures, verification techniques, and evaluation metrics, while pointing to common challenges and open research directions [3]. Other comprehensive surveys on continuous authentication compile supervised and unsupervised learning techniques, discuss context-aware approaches, and catalogue challenges related to scalability, dataset quality, adversarial robustness, and lack of standardized metrics [4]. Parallel work on behavioral biometric continuous authentication for smartphones concentrates predominantly on technical system design and performance, while user attitudes, adoption drivers, and perceived usability of these technologies remain under-explored [1]. In the IPA domain, systematic mappings identify taxonomies, usage domains, security and privacy issues, and under-served user profiles, yet dedicated integration of continuous behavioral biometrics into IPA-centric security architectures has not been thoroughly surveyed [5]. Against this backdrop, a systematic review that jointly examines continuous mobile user authentication via behavioral and multimodal biometrics, its relationship to intelligent personal assistant ecosystems, and its embedding in modern IoT and mobile network contexts can synthesize fragmented insights and clarify future research needs. By drawing on the methodological rigor and gap analyses provided by prior surveys in continuous authentication, behavioral biometrics, and IPA research, the present work positions itself to assess machine learning advances, contextual and wearable integrations, and privacy-preserving designs that shape the next generation of biometric security for mobile and assistant-enriched environments [1,3–5].

While existing surveys address continuous mobile authentication from complementary but partial perspectives—Abuhamad et al. [3] restrict their scope to smartphone sensor modalities without addressing distributed architectures, Ayeswarya and Singh [4] extend coverage to IoT and blockchain but omit federated learning and IPA integration, and Baig and Eskeland [9] provide rigorous security-privacy-usability analysis yet exclude edge computing, federated learning, and agent-based architectures—none jointly examines the

interplay between behavioral biometric modalities, intelligent personal agent pipelines, and privacy-preserving distributed infrastructures. The present review addresses this gap by: (i) systematically positioning IPA components as active authentication agents across device, edge, and cloud tiers; (ii) incorporating federated learning alongside ZKP, cancelable biometrics, and Bloom filters within a unified privacy threat model; and (iii) mapping cross-domain deployment challenges at the intersection of mobile devices, wearables, IoT ecosystems, and assistant-mediated authentication—an integration that prior surveys have identified as needed but left uncharted.

Table 1 provides a structured comparison of the most closely related systematic reviews across twelve analytical dimensions. As the comparison reveals, no prior survey jointly addresses IPA pipelines, federated learning, agent-based architectures, and PRISMA-compliant methodology within a unified framework. The present review addresses this gap by synthesizing 80 studies selected through a transparent PRISMA-based protocol, and by explicitly positioning IPA components as active authentication agents across device, edge, and cloud tiers.

Table 1. Comparison of related systematic reviews on continuous authentication and behavioral biometrics.

Dimension	Abuhamad et al. [3]	Ayeswarya & Singh [4]	Baig & Eskeland [9]	Finnegan et al. [10]	Present Review
Scope/Focus area	Smartphone sensors	CA & user profiling	Security, privacy & usability	Screen-time utility	BBCA, IPA & IoT
Behavioral biometrics	✓	✓	✓	✓	✓
Physiological biometrics	Partial	✓	✓	—	✓
IPA/voice assistant pipelines	—	—	—	—	✓
Federated learning	—	—	—	—	✓
Edge/fog computing	Partial	Partial	—	—	✓
ZKP/cancelable biometrics	—	✓	✓	—	✓
Adversarial attack analysis	Partial	✓	✓	—	✓
Agent-based architecture	—	—	—	—	✓
PRISMA methodology	—	✓	—	✓	✓
Publication years	–2020	–2024	–2021	2007–2022	2017–2025
Focused primary studies	>140	n/a	n/a	122	80 ^a

Note: ✓ = covered systematically; — = absent; Partial = mentioned but not systematically analyzed. BBCA = behavioral biometric continuous authentication; IPA = intelligent personal agent; CA = continuous authentication; ZKP = zero-knowledge proof. ^a Stricter PRISMA inclusion criteria applied; broader scope compensated by thematic depth across seven research areas.

This gap is not merely taxonomic. IPAs already mediate security-sensitive actions, including voice-activated commands and persistent storage of location and speech data, and are documented as vulnerable to remote manipulation of voice activation and to a lack of standardized security and privacy benchmarks [5]. Because no prior survey jointly analyzes behavioral biometric continuous authentication alongside agent-based architectures, existing guidance offers no consolidated account of *where* continuous behavioral verification checkpoints should sit relative to an agent’s decision points across device, edge, and cloud tiers, nor of how privacy-preserving mechanisms such as zero-knowledge proofs, cancelable biometrics, and federated learning interact when deployed jointly rather than in isolation. This matters concretely for threat modeling: curious or semi-malicious cloud-side insiders who can already access location data, IP addresses, and partial biometric templates during agent-mediated processing [9] face a substantially larger attack surface once agent pipelines, sensor streams, and privacy-preserving primitives are combined, yet the present

literature provides no synthesized account of this compounded threat. By bringing these strands together, this review converts a set of independently documented risks into an explicit research agenda for agent-integrated behavioral authentication, rather than treating the synthesis itself as the contribution.

The remainder of this paper is organized as follows: Section 2 presents the literature review across behavioral modalities, machine learning methods, and security threats; Section 3 details the methodology; Section 4 reports results including publication trends and thematic distribution; Section 5 discusses key findings and research gaps; Section 6 addresses ethical considerations; and Section 7 presents conclusions and future research directions.

The deployment context for continuous authentication is further shaped by three interacting factors that are treated in depth later in this review rather than in the Introduction. First, successive mobile network generations (2G–6G) have progressively reshaped the security and resource envelope available to on-device authentication, a trajectory analyzed in detail in Section 2.1. Second, traditional password- and PIN-based mechanisms remain structurally exposed to low entropy, shoulder surfing, and observation attacks, as detailed in Section 2.2.1. Third, phishing, spoofing, and insider threats against biometric templates and machine learning pipelines constitute a persistent attack surface, examined in Section 2.7.1. Consolidating this material outside the Introduction allows the present section to focus exclusively on the motivation, identified research gap, and contribution of this review.

1.1. Research Questions and Objectives

Clarifying the primary research questions is essential for structuring this review around clearly defined gaps and themes already identified in prior work on continuous authentication, behavioral biometrics, and intelligent personal assistants. Existing surveys on continuous authentication explicitly formulate multiple questions concerning accuracy, usability, privacy, deployment costs, temporal stability of behavior, and real-world application domains, and these provide a concrete foundation for the present study. Stated explicitly, the objective of this review is to determine how behavioral biometric continuous authentication can be systematically integrated with intelligent personal agent architectures across device, edge, and cloud tiers, and to synthesize the associated performance, usability, security, and deployment evidence into a single evaluative framework. This objective is distinct from, and builds on, the four most closely related reviews summarized in Table 1: Abuhamad et al. [3] and Ayeswarya and Singh [4] pursue a comparable evidence-synthesis objective for sensor-based and profiling-based continuous authentication respectively, but neither treats IPA integration as a first-class object of analysis; Baig and Eskeland [9] share this review's emphasis on security and privacy trade-offs but stop short of a PRISMA-compliant protocol; and Finnegan et al. [10] address a narrower screen-time-oriented objective. The five research questions below (RQ1–RQ5) operationalize this objective into the specific dimensions synthesized in Sections 3.5, 3.6 and 4.

RQ1 (Performance and modality choice). A first overarching question concerns performance and modality choice: how well do current continuous authentication systems operate on mobile and related platforms, and which biometric categories provide the most effective basis for continuous user verification. Prior work separates physiological biometrics, such as fingerprint and iris, from behavioral modalities, such as keystroke dynamics and gait, and asks how these types individually and jointly contribute to the effectiveness of continuous authentication systems. Closely related is the question of whether multimodal combinations of physiological and behavioral signals, including fusion with contextual information, can deliver more robust and versatile authentication frameworks than unimodal

schemes in realistic use scenarios. Addressing these issues requires comparing performance metrics such as False Acceptance Rate, False Rejection Rate, and derived measures across modalities and fusion strategies, as well as evaluating how far machine learning advances have improved error rates in contemporary systems. At the same time, reported variability in per-user outcomes and maneuver-dependent discriminability suggests modality selection and fusion strategies must explicitly accommodate heterogeneity in intra-user stability and environmental conditions rather than assuming uniform gains.

RQ2 (Usability and user acceptance). A second group of questions targets usability and user acceptance. Research explicitly asks whether continuous authentication approaches are easy for people to use and what design measures can render them more user-friendly. Within this domain, Continuous User Biometric Authentication Systems are described as facing persistent trade-offs between usability and security, particularly in relation to FRR and FAR, and it is therefore necessary to investigate which configurations and feedback mechanisms can achieve acceptable balances for mobile users in everyday situations [4]. Findings from behavioral mobile authentication indicate that user attitudes, perceived vulnerability, and self-efficacy condition intention to adopt behavioral biometric continuous authentication, which motivates further questions about how psychological constructs and risk perceptions mediate the willingness to rely on continuous behavioral monitoring in practice [1]. The review consequently asks which usability barriers and perception-driven inhibitors most strongly limit adoption, and how they interact with technical properties such as power consumption on sensor-rich smartphones [1,4].

RQ3 (Security, privacy, and threat models). A third set of questions focuses on security, privacy, and threat models. Existing surveys explicitly ask whether continuous authentication systems keep user information safe and how they balance security guarantees with privacy protection. This leads to more detailed questions concerning susceptibility of behavioral and context-aware modalities to spoofing, generative attacks, and adversarial perturbations, as well as the impact of threat actors such as curious or semi-malicious cloud insiders who may gain access to biometric templates and contextual streams during processing [4,9]. Since continuous authentication often implies continuous sensing and data transmission, the review also examines which privacy-preserving mechanisms, including cancelable biometrics, Bloom filter-based template protection, and zero-knowledge proof protocols, are most appropriate for mitigating insider misuse and template reconstruction risks in mobile and cloud-assisted settings.

RQ4 (Deployment constraints, costs, and temporal dynamics). A fourth cluster of questions addresses deployment constraints, costs, and temporal dynamics. Prior work formulates explicit questions about what factors must be considered when setting up continuous authentication systems and how expensive such deployments are in terms of computation, memory, communication, and energy overheads. Smartphone-focused analyses highlight power consumption of sensors and continuous processing as a major difficulty, which in turn motivates systematic inquiry into how resource constraints on mobile and IoT devices limit feasible modality choices and model complexity. Further, continuous-authentication surveys ask how user behavior changes over time and how systems can adapt to these changes. This review therefore investigates which update strategies, re-enrollment mechanisms, and temporal modeling approaches have been proposed to maintain performance in the face of behavioral drift without incurring excessive retraining or user burden.

RQ5 (Application domains and IPA integration). Finally, guided by questions about real-world usage, prior surveys explicitly seek to determine where continuous authentication has been successfully deployed, which domains remain underexplored, and what examples exist of systems operating in practice [4]. Synthesizing these strands, the present

review asks which application areas at the intersection of mobile devices, IoT, and IPA ecosystems are most promising for behavioral continuous authentication, how security and privacy issues within IPA usage constrain such integration, and what open questions remain regarding proactivity and personalization of assistant-mediated authentication [4,5].

1.2. Contributions

This review makes four main contributions. First, it synthesizes 80 studies (2017–2025), selected via a PRISMA-compliant protocol, into the first systematic account that jointly covers behavioral biometric continuous authentication on mobile devices *and* its integration with intelligent personal agents, wearables, and IoT/edge infrastructures; prior surveys treat continuous authentication and intelligent personal assistants as largely separate research strands [4,5]. Second, it provides a structured cross-modality comparison, organized around the taxonomy in Figure 1, that links each behavioral modality (keystroke dynamics, touch/swipe, gait/motion, contextual data) to reported performance metrics, security properties, and energy/latency trade-offs on resource-constrained devices, consolidating evidence that is otherwise dispersed across modality-specific surveys [4,11]. Third, it evaluates privacy-preserving mechanisms, including cancelable biometrics, Bloom filter encodings, zero-knowledge proof protocols, federated learning, and blockchain-based identity management, against concrete deployment constraints (on-device versus cloud-assisted processing, agent–sensor–TEE interaction) rather than discussing them in the abstract; for example, agent-driven facial capture for continuous verification in online examinations illustrates how such mechanisms trade-off against the need for extra sensors and centralized template storage [12]. Fourth, it identifies and substantiates specific, underexplored research gaps from the synthesized evidence: the absence of standardized adversarial benchmarks for behavioral modalities, the lack of end-to-end pipeline evaluations under simultaneous adversarial and privacy threat models, and limited user-centered evidence on consent and acceptance of privacy-preserving mechanisms under frameworks such as GDPR. Methodologically, the review documents its search strategy, screening process, and quality assessment criteria in full (Section 3.5 and Supplementary Materials) to support replication by future systematic reviews at the intersection of behavioral biometrics, mobile security, and intelligent agents. Beyond the narrative synthesis itself, this review produces three reusable artifacts that go beyond a conventional literature summary: (i) the four-factor, four-modality taxonomy in Figure 1, which has not previously been formalized for this specific intersection of behavioral biometrics and IPA architectures; (ii) the three-tier device–edge–cloud reference architecture in Figure 2, which explicitly positions TEE-based template protection, cancelable biometrics, and federated aggregation at specific architectural boundaries rather than discussing them as independent mechanisms; and (iii) the consolidated attack-defense mapping presented in Section 2.7.1, which links specific reported attack success rates (e.g., 5.8–43.66% FAR increase under treadmill-assisted gait imitation; 28–84% EER increase under keystroke generative attacks) to specific defense mechanisms at a level of granularity not present in any single primary study reviewed here. These artifacts are intended as direct inputs for subsequent empirical work (Section 5.2), rather than as restatements of existing survey content.

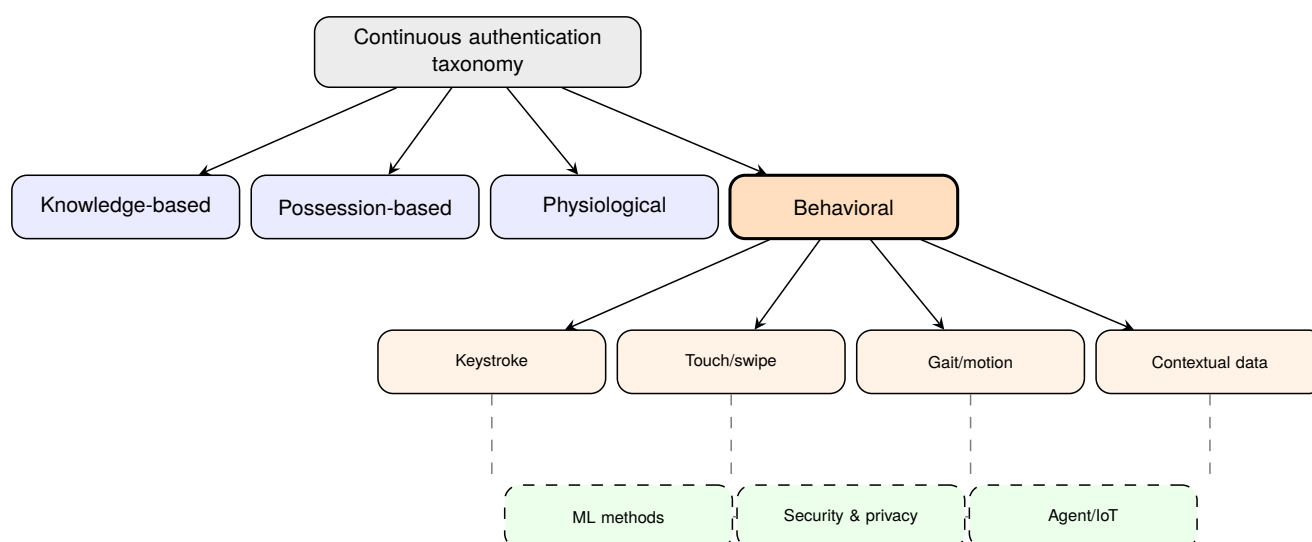


Figure 1. Taxonomy of continuous authentication factors and modalities adopted in this review, showing the four top-level authentication factors, the four behavioral modalities analyzed in detail, and the three cross-cutting layers (machine learning methods, security and privacy, and agent/IoT integration) that recur across modality-specific discussions.

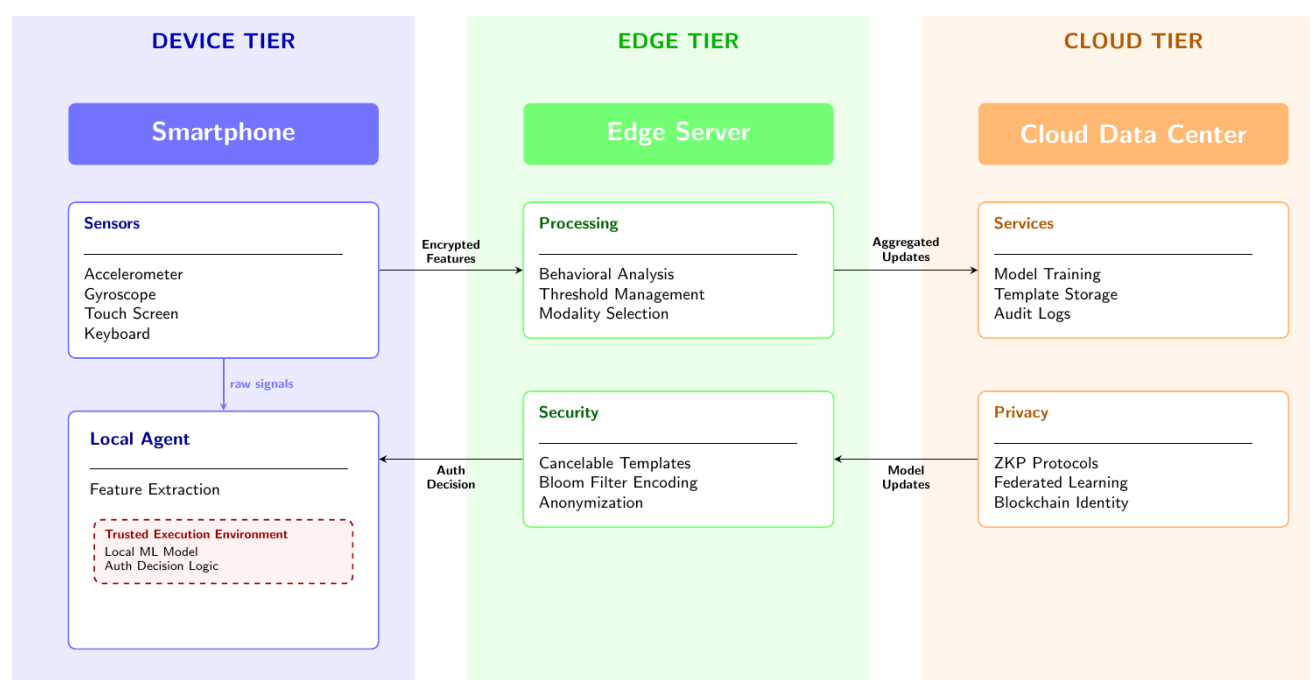


Figure 2. Three-tier architecture of an intelligent continuous authentication system spanning device, edge, and cloud layers. Within the device tier, raw sensor signals are passed to the local agent's feature extraction stage, while the local ML model and authentication decision logic operate inside a Trusted Execution Environment (TEE) to protect behavioral templates and decision logic from the host operating system.

2. Literature Review

Behavioral biometrics aim to authenticate users from how they act rather than what they know or their static physical traits. Typical modalities include gait, gestures, tapping patterns, and daily activities captured by inertial sensors or WiFi channel state information. These behavioral characteristics are often difficult to consciously reproduce, which can reduce the success of imitation and random attacks compared with knowledge-based or tra-

ditional biometric factors. Moreover, multi-sensor fusion, for instance, combining touch and accelerometer signals, can raise the security level under observation attacks by increasing the dimensionality and distinctiveness of behavioral signatures. Behavior-based schemes can also operate transparently in the background, which improves usability by avoiding explicit authentication steps and supports continuous or recurrent verification [13,14].

Several works focus specifically on gait and daily activity based authentication on resource-constrained mobile or wearable devices. Lightweight models that exploit accelerometer or related inertial data can meet design requirements such as low computational overhead, real-time classification, and high identification accuracy with only a small number of embedded sensors. Approaches that merge temporal and spatial features enable spatiotemporal extraction from walking and computational activities, providing richer behavioral representations. Other architectures propose logical DNA-style profiles that integrate multiple behavioral factors to enable continuous authentication beyond a single activity. Collectively, these techniques indicate that behavioral biometrics can simultaneously strengthen security and preserve user convenience on mobile platforms [13,14].

The threat landscape for biometric and behavior-based systems includes several classes of adversaries and attack vectors. For mobile authentication in general, brute force attacks attempt large numbers of identity related values, which is computationally and temporally expensive, while guessing attacks opportunistically exploit line-of-sight observation of entry actions such as typing a password. Observation attacks can directly reveal knowledge-based secrets and also threaten biometrics when the adversary captures sufficient information about user interactions or movements. Impersonation attacks then use observed or inferred identity information, such as PINs or gesture patterns, to assume the victim's identity. In behavioral systems, random attacks by strangers, who lack prior knowledge of the victim, typically have near zero success, whereas imitation attacks, even by motivated observers, show very low success because fine-grained behavioral patterns are hard to replicate faithfully [13,15].

Biometric research explicitly distinguishes between insider and stranger adversaries. Insider adversaries are familiar with the victim, can observe biometric behaviors in proximity, and may know protocol details and user identifiers. Such insiders can mount active, skilled forgery attacks, for example, by visually observing in-air signing gestures and practicing them before attempting impersonation. In contrast, stranger adversaries lack detailed prior knowledge and generally mount random attacks based only on generic understanding of the biometric modality. This adversary taxonomy is crucial for evaluating behavioral schemes, whose robustness can differ sharply between random, observation based, and skilled forgery scenarios [13,15].

With the increasing deployment of machine learning in biometric pipelines, adversarial attacks against the underlying models have become a central concern. In the behavioral domain, white box and black box adversarial strategies have been analyzed for classifiers such as *k* Nearest Neighbors, Random Forests, and one-dimensional convolutional neural networks. In white box settings the attacker is assumed to know model architecture, weights, and feature space, which allows direct gradient based perturbation of inputs, whereas black box attackers only query the system and must infer effective perturbations without internal access. For practical mobile or wearable authentication, it is often unrealistic to assume that adversaries have full white box access, so black-box-transfer-based attacks may be the more relevant threat. Nonetheless, the existence of both attack classes underscores the need for robust behavioral models and for careful threat modelling in system design [14,16].

Federated learning introduces additional opportunities and challenges for behavioral biometrics and other security critical applications. In smart eHealth scenarios, federated

learning is used to verify system attributes in the cloud while patient data are collected by medical entities and transmitted through a secure gateway. Such architectures aim to achieve mutual authentication, data privacy, anonymity, session unlinkability, and non-traceability between devices and access gateways. Federated setups thus hold promise for privacy-preserving training of behavioral models, since raw sensor data can remain on device while only model updates are aggregated. However, they also introduce communication and resource constraints, which can be mitigated by techniques such as gradient compression that sparsify weight matrices and reduce transmission overhead. Careful integration of behavioral biometrics with federated learning is therefore a key direction for privacy-aware, attack-resilient authentication in emerging mobile and eHealth systems [8,14,17].

Despite these advances, several open challenges remain. Standardized evaluation frameworks that cover insider versus stranger adversaries, white box versus black box attacks, and federated threat models are still lacking. The heterogeneity of sensors, variable sampling rates, and the need for energy-efficient on-device processing complicate reproducibility and cross-study comparison. Moreover, emerging studies indicate that adversarial robustness techniques developed for image or speech domains do not always transfer directly to time-series behavioral signals, necessitating domain-specific defenses and adaptive detection strategies. Future work should therefore prioritize (i) unified benchmarks for behavioral modalities under realistic adversarial and federated conditions, (ii) lightweight certified or empirically validated defenses suitable for mobile devices, and (iii) privacy-preserving personalization methods that resist poisoning and inference while maintaining usability.

In summary, behavioral biometrics offer promising avenues for transparent, continuous authentication with improved resistance to casual and random attacks, while federated learning provides mechanisms for privacy-preserving model training. However, the combination of behavioral modalities, resource constraints, and federated architectures creates a complex security landscape in which adversarial attacks, from transfer-based black box perturbations to client-side poisoning and backdoors, must be explicitly considered. Recent literature has begun to address these intersections by proposing multimodal schemes, federated defensive measures, and adversarial analyses specific to behavioral signals, yet coordinated benchmarks and deployable, lightweight defenses remain important priorities for the community.

2.1. Mobile Network Evolution and Resource Constraints for Continuous Authentication

Mobile communication systems have progressed through successive generations that differ markedly in data rates, service capabilities, and security properties. Second-generation (2G) networks introduced digital transmission and basic privacy mechanisms such as Temporary Mobile Subscriber Identity (TMSI) and radio path encryption, yet end-to-end protection remained weak when only parts of the communication channel were encrypted, leaving remaining segments exposed to eavesdropping attacks.

Third-generation (3G) systems, launched around 2000 with data rates up to 2 Mbps, enabled services such as TV streaming, internet browsing, and video streaming that had not been feasible on earlier mobile communication systems. To mitigate known 2G vulnerabilities, 3G adopted two-way authentication and an Authentication and Key Agreement (AKA) protocol, supported by a comprehensive Third-Generation Partnership Project (3GPP) access control security system covering air interface security and user authentication on both sender and receiver sides. Nevertheless, 3G environments still face threats related to Internet Protocol vulnerabilities, attacks on communication channels between end devices and home networks, and wireless interface threats including integrity violations, unau-

thorized data access, denial of service, and unauthorized service access, alongside AKA privacy issues such as sniffing of user identities.

Fourth-generation (4G) networks, commercialized from 2009 onwards, increased downlink transmission up to 1 Gbit/s and uplink up to 500 Mbit/s, and improved spectrum efficiency and latency to support demanding applications such as High-Definition Television and Digital Video Broadcasting. Architecturally, 4G systems integrate IP core networks, backbones, diverse access networks, and intelligent mobile terminals, which broadens the attack surface. Primary 4G security problems arise from threats to wireless radio communication, including tampering, eavesdropping, data alteration, and weaknesses in network authentication. Additional issues stem from increased indirect interaction between users and mobile terminals and from enhanced storage and computing capabilities on devices that enable tampering with hardware platforms, malware infections, and operating-system-level attacks. Medium Access Control layer vulnerabilities such as eavesdropping, replay, data integrity attacks, unauthorized access, and location tracking through MAC protocols further complicate the security posture of 4G deployments.

Fifth-generation (5G) architectures introduce higher data speeds and complex systems that aim for high security, while simultaneously connecting a rapidly growing number of heterogeneous devices and access methods. The architecture spans access, backhaul, and core networks, and frequent handovers across different access technologies and device types increase the probability of attacks during mobility. Backhaul segments, implemented via microwave, wireless, satellite, or traditional links, tend to present fewer privacy risks than access networks, yet security concerns can still be propagated toward the core network. At the same time, integration of Software Defined Networking and Network Functions Virtualization into the control and data planes yields highly dynamic infrastructures in which virtualized services may migrate across resources, creating additional threats and vulnerabilities. Large multiple-input multiple-output configurations are deployed to conceal active and passive eavesdropping, while SDN implementations via OpenFlow can be undermined by rogue applications or activities. NFV migration dynamics and the open nature of 5G platforms raise serious privacy issues, since user information can be easily exposed, and these concerns are expected to intensify as 5G usage grows. Enhanced Mobile Broadband services raise signaling loads and denial of service risks, motivating lightweight authentication, group-based AKA mechanisms, and other security measures that preserve service continuity and quality while managing latency introduced by security processes [6]. These network-layer security mechanisms have direct architectural implications for behavioral biometric continuous authentication. The same edge-distributed structure that 5G uses to reduce handover latency also offers a natural placement for the device–edge–cloud authentication pipeline introduced in Section 2.5.2: behavioral feature extraction can run on-device, lightweight verification at the edge, and only aggregated, privacy-protected updates need to reach the core network, reducing exposure of raw sensor streams to the same NFV-orchestrated infrastructure whose dynamic service migration was identified above as a privacy risk. Latency-sensitive V2X deployments illustrate this trade-off concretely: a blockchain-based identity scheme for 6G-enabled V2X communications assigns each vehicle a digital identity on-chain specifically to meet the sub-millisecond authentication budgets that URLLC service classes require, while still preserving an auditable, tamper-resistant authentication trail [18].

Each of these three mechanisms creates a distinct, technically separable implication for behavioral authentication that the general discussion above does not fully draw out. SDN separates the control plane from the data plane, giving a centralized controller programmatic visibility over traffic flows; this enables an architecture in which a continuous authentication decision (e.g., a sustained anomaly score for a given session) can trigger

a network-level policy change, such as throttling or quarantining traffic from a device, enacted directly through SDN flow rules rather than through an application-level access control list. This shifts part of the authentication enforcement boundary from the device or application layer into the network fabric itself, but it also means that an attacker who compromises the SDN controller gains the ability to override authentication driven traffic policies for many devices simultaneously, concentrating risk that was previously distributed across individual endpoints. NFV, in turn, allows authentication logic, such as a behavioral feature extraction or scoring service, to be deployed as a virtualized network function that can be instantiated, scaled, and migrated independently of the underlying hardware; this elasticity is attractive for handling variable continuous authentication load across many simultaneous mobile sessions, but each migration event is also a moment at which behavioral templates or model parameters in transit could be exposed to the host infrastructure, echoing the privacy concerns already raised for NFV service migration in Section 2.1. URLLC, finally, defines a 5G service class targeting round trip latencies below one millisecond with very high reliability, intended for safety-critical use cases; while this budget is well suited to transmitting a short authentication decision or revocation signal, it is poorly matched to the inference latency of most behavioral biometric pipelines reviewed in this work, particularly deep learning-based keystroke, touch, and gait models (Section 2.6.2), whose feature extraction and classification steps routinely take tens to hundreds of milliseconds on mobile hardware. This mismatch implies that URLLC, as currently specified, can support fast propagation of an already computed authentication decision but cannot by itself make behavioral biometric inference real time, leaving the computation step, not the network step, as the binding latency constraint for continuous authentication in 5G and 6G deployments.

Looking ahead, proposals for future sixth-generation (6G) systems anticipate even greater network capacity and more stringent application requirements than those of established 5G networks. Increasing device density and service diversity will further amplify signaling loads and the impact of security mechanisms on ultra reliable low latency communications. Security in this context is framed as a high requirement for ensuring availability and continuity of services and resources, with explicit recognition that latency penalties stemming from security processes must be carefully addressed. As networks become more dynamic due to NFV, SDN, and cloud technologies, research emphasizes the necessity of effective security solutions that can operate at scale without undermining performance or reliability [6]. Beyond-5G (B5G) proposals already anticipate this convergence of network and authentication concerns in domains adjacent to mobile continuous authentication. In B5G-enabled smart grid deployments, digital twin architectures combine intelligent authorization with the higher device density and stricter latency requirements anticipated for 6G, suggesting that continuous, context-aware authentication mechanisms developed for mobile users may need to be re-evaluated for non-human, machine-type endpoints operating under the same network constraints [19]. For mobile behavioral biometric authentication specifically, this implies that future 6G-era architectures will need to budget authentication latency, not only raw data throughput, as an explicit network resource, since the device–edge–cloud partitioning decisions discussed for IPA-mediated continuous authentication are directly constrained by the same URLLC guarantees that 6G research currently frames primarily in terms of generic service continuity rather than biometric decision latency.

In parallel with network evolution, modern smart devices have accumulated numerous embedded sensors and greater computational resources, which has direct implications for continuous authentication. Smartphones typically incorporate accelerometers, gyroscopes, magnetometers, light sensors, proximity sensors, and digital compasses, enabling

rich monitoring of user interactions and context. However, continuous sensor utilization consumes substantial battery power, and sensory data collection can significantly increase energy usage, particularly for Continuous User Biometric Authentication systems that require uninterrupted processing and sensor activity throughout active sessions. These constraints illustrate how the hardware evolution that enables sophisticated behavioral biometrics simultaneously introduces power management challenges for continuous authentication on mobile platforms [4].

One pragmatic counterproposal appears in the Trusted Distributed Artificial Intelligence approach, which argues for on-board trusted agents and Trusted Execution Environment-based networking that shift analytic workloads toward edge units and thereby may reduce unnecessary sensor transmission and central processing overhead. Such distributed and hybrid architectures also permit locally enforced trust metrics and selective sharing policies to limit privacy exposure, although their practical value will depend on careful calibration of trust thresholds and the overhead of the monitoring mechanisms [20]. Users frequently struggle to weigh the benefits of uninterrupted biometric monitoring against the privacy and security costs that such monitoring incurs; a modest way to address this is to employ assistance that models the utility–privacy trade-off and offers actionable recommendations. The Online Information-Sharing Assistance framework proposes a game theoretic representation of those trade-offs and demonstrates how AI agents can improve user outcomes by balancing benefits, costs, and privacy losses, even while cautioning that computational complexity and personalization remain practical hurdles for real deployments [21].

Growing digital connectivity reshapes everyday practices by embedding authentication into routine interactions with smartphones and connected systems. Mobile devices now mediate banking, communication, storage of personal photos, and access to numerous online services, which intensifies the consequences of unauthorized access for users in daily life [22]. Traditional, one-time mechanisms such as passwords and PINs have been shown to be inadequate for protecting smartphones once a session is established, since loss or theft during an active session can expose sensitive data. Continuous authentication, particularly behavioral biometric continuous authentication, directly targets this vulnerability by monitoring behavior across the session rather than only at login [1]. As a result, security enforcement increasingly accompanies users throughout routine activities, rather than appearing as a single discrete step. Smartphones incorporate multiple embedded sensors, including accelerometers, gyroscopes, proximity sensors, light sensors, barometers, and digital compasses, and these capabilities enable the capture of fine-grained activity and motion patterns for behavior-based authentication. Frameworks such as DeepAuthen demonstrate that continuous authentication can be realized by exploiting such multimodal sensing to infer user activities without requiring explicit interaction, enabling validation while users perform ordinary tasks [22]. Continuous User Biometric Authentication Systems, however, require ongoing sensor activity and processing throughout active sessions, which raises practical concerns for users regarding battery life and device responsiveness. Modern smartphones already face battery life as a major difficulty, and sensor usage together with sensory data collection consumes substantial power, so continuous monitoring alters how users must think about charging habits and device configuration during daily use [4,22].

From a usability perspective, continuous authentication introduces a persistent trade-off that directly affects user comfort and acceptance. Continuous User Biometric Authentication Systems must simultaneously minimize False Acceptance Rate and False Rejection Rate, since high FAR undermines security while elevated FRR leads to frequent legitimate user denials and frustration [4]. Research on user verification explains that FAR and FRR can be tuned by adjusting probability thresholds, and that Equal Error Rate is often used

as a summary measure of biometric performance, linking technical design choices to the practical frequency of disruptions a user experiences [14]. Empirical work on mobile keystroke biometrics using Transformer architectures has shown that free-text verification may reach surprisingly low EERs with modest enrollment data, which in turn suggests that behavior-based continuous authentication could be practically effective if algorithmic design and enrollment strategies are carefully balanced [23]. The study further notes trade-offs between per-user threshold tuning and global deployment, a tension that appears to mirror concerns about computational cost, privacy, and the burden of continuous sensing discussed above. Studies further show that smartphone users are interested in continuous authentication but express a preference for methods with lower resource consumption and voice concerns about privacy when behavioral data are leveraged for security purposes [9]. These considerations reveal that authentication is no longer a purely background security function; instead, it shapes perceived convenience, trust, and willingness to adopt mobile services in everyday contexts.

Privacy and ethical implications also permeate daily life when authentication relies on continuous sensing and behavioral profiling. Continuous authentication systems may collect and store biometric and contextual information, potentially monitoring user activities in ways that are not always transparent to end users [4]. Reviews of Intelligent Personal Assistants identify major challenges around usability, security, and privacy that influence technology confidence and adoption, with many people restricting their assistant usage to simple, low sensitivity tasks such as listening to music or checking the weather, thereby avoiding scenarios that require sharing more private data [5]. Systematic analysis of continuous authentication further documents that security and privacy remain contentious, even when systems appear usable, and stresses that stakeholders must explicitly address these concerns before deploying solutions in specific domains. For users, this means that decisions about which services to enable, how much data to share, and whether to trust always-on assistants or behavioral monitoring tools increasingly become part of everyday digital hygiene.

Future directions in continuous authentication and privacy-preserving mechanisms will likely continue to influence user lifestyles as emerging methods are transferred into practical applications. Proposed strategies such as zero-knowledge-proof-based protocols, cancelable biometrics, and template protection using Bloom filters are explicitly designed to safeguard biometric templates and limit exposure of sensitive authentication data, aiming to reduce privacy risks while maintaining security. At the same time, research advocates interdisciplinary efforts that integrate psychological insights into emotional states and behavior, together with user surveys, to design solutions that are more acceptable and better aligned with how individuals actually interact with devices in real settings [4]. As continuous authentication expands into mobile, IoT, and assistant-based ecosystems described in Section 2.1, these technical and ethical developments will continue to shape how users manage identity, security, and privacy as ordinary components of their digital routines [4,5,22]. Some implementation studies suggest pragmatic mitigations for energy and computation costs by aggregating touch and keystroke features locally into sliding windows and transmitting averaged packets to a server that holds per-user interaction models, an approach that can cut on-device processing but does shift storage and inference to a central service [24]. Such architectures may also generate simple, quartile-based visual explanations for users and have reported high accuracy in small trials, yet they appear to trade continuous sensor use and centralized data retention for improved detection rates, a trade-off that may influence user acceptance in practice.

Before examining each authentication factor in detail, Figure 1 summarizes the taxonomy adopted throughout this review. Continuous authentication mechanisms are orga-

nized into four top-level factors: knowledge-based, possession-based, physiological, and behavioral. Behavioral biometrics, which form the primary focus of this survey, are further decomposed into four modalities discussed in the sections that follow: keystroke dynamics, touch/swipe gestures, gait and motion patterns, and contextual data (the latter detailed in Section 2.4.4). These modalities are not analyzed in isolation; three cross-cutting layers, namely machine learning methods, security and privacy mechanisms, and agent/IoT integration, recur throughout the modality-specific discussions in Section 2.2 and beyond, and are treated explicitly in later sections of this review.

2.2. Classical Authentication Factors and Their Limitations

2.2.1. Knowledge-Based Factors

Knowledge-based authentication relies on secrets known only to the legitimate user and the system, typically implemented as textual PINs, alphanumeric passwords, lock patterns, and graphical or color-based secrets on mobile devices. Two broad categories can be distinguished: text-based secrets, which use character strings as the credential, and graphical secrets, which encode the secret in images or gestures rather than in text. Text-based schemes remain the most popular and widely deployed methods for mobile access control, with one report cited by Wang et al. stating that PINs and passwords are chosen by 66% of consumers as their preferred mobile authentication mechanism. These schemes usually enforce minimum lengths, such as four or six digits for PINs and at least eight characters for passwords, in order to resist random guessing and brute force attacks through increased combinatorial space [13]. Table 1 provides a comparative overview of major authentication approaches across key security, usability, and operational dimensions. The Low/Medium/High/Very High labels in this table are qualitative, ordinal summary judgments rather than scores computed by this review's own quality-assessment rubric (Section 3.5); they synthesize directional trends reported consistently across the cited literature (e.g., "password entropy is low relative to biometric discriminability") rather than a numeric average of any single reported metric. Where the reviewed literature reports a quantitative range for a given cell (for example, PIN entropy of 13–20 bits versus password entropy of 20–78+ bits, discussed below), that range is given in the surrounding text rather than in the table itself, and the qualitative label should be read as a relative ranking across the four method categories rather than as an absolute security guarantee. Ratings are therefore best interpreted column-wise (comparing categories against one another) rather than row-wise as fixed thresholds. Table 2 summarizes these criteria side by side for password/PIN, physiological biometric, behavioral biometric, and multimodal/continuous authentication approaches.

Table 2. Comparison of authentication methods.

Criteria	Password/PIN	Physiological Biometrics	Behavioral Biometrics	Multimodal/Continuous
Security Level	Low	High	Medium–High	Very High
Usability	High	Medium	High (passive)	Medium
Continuous Auth.	No	No (point-of-entry)	Yes	Yes
Shoulder Surfing Risk	Very High	Low	Very Low	Very Low
Spoofing Resistance	Low	Medium	Medium	High
Privacy Risk	Low	High	Medium	Medium
Energy Consumption	Very Low	Low	Medium	High
Revocability	Yes	No	Partial	Partial
FAR (typical)	High	Low	Medium	Low
Implementation Cost	Very Low	Medium–High	Low	Medium

Note: FAR—False Acceptance Rate. Ratings reflect general trends reported in the literature [2,4,9,13], synthesized qualitatively rather than computed from a single quantitative source; see the discussion following this table for the specific entropy and vulnerability figures underlying each rating.

Analysis of security usability trade-offs shows that PINs offer relatively low entropy, around 13 to 20 bits, whereas passwords can reach higher entropy ranges between 20 and 78 or more bits under stringent composition policies. This higher secrecy, however, comes with reduced usability, since more complex passwords are harder to remember and slower to input. Modern approaches such as two-factor authentication, biometrics, and adaptive schemes are explicitly framed as attempts to rebalance this trade-off between secrecy and convenience in knowledge-based authentication [2]. Empirical work on user choice behavior further reveals that people frequently select weak secrets based on birthdays, addresses, phone numbers, repeated digits, or sequential patterns, which are easy to recall but highly susceptible to guessing and dictionary attacks. Surveys report that such simple PINs and passwords can result in one unauthorized bank account access for every 11 to 18 stolen wallets, illustrating the concrete impact of predictable secret selection [13].

Observation and side channel threats compound these structural weaknesses. Shoulder surfing and smudge attacks exploit the visibility of input on touchscreens and oily residue patterns left on glass surfaces, enabling reconstruction of PINs or patterns after brief physical access to the device. Mobile attack overviews explicitly list shoulder attacks, brute force, dictionary, replay, phishing, key logging, guessing, and smudge attacks as prominent threats against PIN and password based mechanisms, and reviews of mobile phone authentication describe traditional passwords and PINs as having numerous disadvantages that motivate investigation of biometric alternatives. Two-factor mobile money schemes using only a SIM and a PIN are likewise characterized as vulnerable and in need of stronger mechanisms, which underscores the fragility of pure knowledge-based designs in adversarial mobile contexts [2]. Knowledge leakage through social engineering and side channels is recognized as a fundamental issue; any adversary who can observe, record, or infer the secret via display contents, timing, motion, or wireless emanations can bypass the authentication barrier [13].

Graphical and gaze-based schemes seek to mitigate some of these problems but introduce their own limitations. Recognition, pure recall, and cued recall graphical passwords leverage human aptitude for remembering images and are reported to be more memorable than alphanumeric passwords, yet they remain exposed to phishing, screenshot based attacks, and predictable “hot spot” selection on images. For recognition-based methods, attackers may capture screenshots or mislead users into interacting with counterfeit image sets, while cued recall schemes face guessable click regions and time consuming scanning over many pictures. Studies of graphical passwords explicitly note that users often pick images or regions correlated with demographics such as race and gender, which further reduces effective entropy and makes secrets easier to infer [2,13]. Entropy analyses classify graphical authentication as providing moderate to high theoretical entropy, in the range of 25 to 40 bits, but emphasize that practical entropy is degraded by user behavior and attacks such as shoulder surfing and smudges, again highlighting the tension between added complexity for security and the resulting usability burden.

Color-based schemes represent another branch of knowledge-based methods, replacing or augmenting textual or image selection with color choices. Reported systems include two-step color codes, hybrids combining faces, points, stories, and colors, and variants like ColorSnakes that extend traditional PINs with color attributes. Although such approaches can offer more personal and meaningful passwords and are described as less directly linked to user information than text-based secrets, they may confuse users, lengthen login times, and exclude color blind participants. Entropy assessments actually rate color-based authentication as providing lower entropy, on the order of 5 to 10 bits, while still being vulnerable to replay and shoulder surfing, which positions these schemes as user-friendly but rela-

tively weak replacements rather than robust successors to conventional knowledge-based factors [2].

Research on future directions for knowledge-based authentication on mobile devices stresses several open problems. First, leakage of PINs, passwords, and lock patterns via shoulder surfing or side channels remains unresolved; proposed mitigations such as randomized keypads can resist some attacks but slow down input because users must visually relocate each key for every login. Second, recognition based graphical passwords need significantly larger password spaces and faster recognition times to become practical. Third, integrating knowledge-based secrets with biometrics and ownership factors at a single input step is proposed as a promising avenue, since integrated multi-dimensional metrics collected simultaneously are harder to forge even when an attacker observes the user's inputs, while also reducing total user effort compared with sequential multi-factor procedures [13]. These findings collectively portray knowledge-based factors as highly usable yet structurally vulnerable components that increasingly require reinforcement through behavioral and multimodal mechanisms in contemporary mobile security architectures [2,13].

Beyond these device-level weaknesses, resource-constrained networked devices have also been addressed through lightweight software and hardware authentication primitives, such as PKI variants, ECC, and Physical Unclonable Functions (PUFs), which may offer acceptable trade-offs between security and device-level constraints [25]. Decentralized, blockchain-backed authentication schemes provide a complementary direction: one design eliminates a central key generation authority, leverages a pairing-free digital signature to cut computational cost, and supports anonymous yet traceable identities with batch verification to improve throughput, which may be appealing for constrained mobile deployments [26]. Such architectural choices can, however, shift complexity into system management and recovery, so they require careful, user-centered evaluation to confirm a genuine reduction in reliance on low-entropy secrets rather than a transfer of burden to administrators or end users.

2.2.2. Possession-Based Factors

Tokens, smart cards, SIMs, and device identifiers exemplify possession-based credentials that are frequently combined with knowledge and biometric factors on mobile platforms. Reviews of user authentication on mobile devices classify ownership-based mechanisms as a distinct metric alongside knowledge and biometrics, and describe ownership factors such as software tokens or hardware identifiers, including MAC addresses, as natural building blocks for two-factor schemes [13]. In mobile money environments, one prominent design uses a Subscriber Identity Module together with a personal identification number for user verification, explicitly relying on the physical SIM as a possession factor and the PIN as a knowledge factor [2]. Similar constructions appear in mobile phone authentication surveys that distinguish between “knowledge + ownership” and “biometrics + ownership” combinations in their taxonomy of approaches, positioning ownership information as an easily fusable component with relatively low standalone security [13].

Security analyses raise important limitations when possession elements operate alone or are weakly bound to user identity. The review of two-factor mobile money schemes reports that existing deployments which only employ a SIM plus PIN remain vulnerable and require security enhancements, despite formally involving two distinct factors [2]. In the broader taxonomy of mobile authentication, it is noted that ownership-based factors provide limited protection when used as a single metric, since adversaries may forge software or hardware tokens and simultaneously supply stolen knowledge-based or biometric information to pass verification. This assessment reflects a core issue: device-bound credentials such as identifiers and tokens can often be cloned or spoofed without requiring

direct interaction with the legitimate user, which undermines their independent security contribution. At the same time, these ownership credentials exhibit attractive usability properties that motivate their widespread integration into multi-factor schemes. Surveys explain that ownership factors are easy to combine with knowledge or biometric metrics “without decreasing the convenience of the authentication process,” because they typically operate transparently once provisioned on the device. For instance, adding a software token derived from device identifiers to a password-based login does not usually demand extra user actions at each authentication attempt. This convenience stands in contrast to traditional multi-factor workflows that require users to sequentially enter passwords, present biometrics, and handle external tokens, which can be tedious when executed repeatedly on smartphones.

Recent work advocates moving beyond naive concatenation of possession and other factors toward integrated, multi-dimensional authentication metrics collected simultaneously. One survey proposes approaches in which knowledge, biometrics, and ownership signals are acquired in a single step via embedded sensors, so that correlations among these dimensions become hard to forge even under observation. Prior implementations that integrate knowledge-based and biometrics based factors illustrate how combining information from distinct physical domains can embed unique identity characteristics of a user; the same line of reasoning extends to blending ownership attributes with these domains. According to this analysis, joint capture of multiple metrics simultaneously both increases security strength and reduces user effort compared to schemes where factors are provided separately.

Despite this potential, several open research directions around possession elements are explicitly identified. First, preventing leakage and replay of authentication input data, including ownership related signals, remains largely unsolved. Empirical analyses of mobile applications indicate that attackers can exploit requested permissions and app packaging to harvest or spoof device-resident tokens and identifiers; studies using static and dynamic analysis show how permissions and system-call traces may be abused to enable malicious behavior that compromises ownership credentials [27]. This suggests that transparent, on-device possession factors may widen the attack surface unless they are bound more tightly to user identity and monitored at runtime. Second, the problem of differentiating genuine user input from replayed sensor data is particularly critical for environments like voice assistant systems and the Internet of Things, where attackers might record and replay user actions combined with device-specific identifiers to spoof authentication. Existing liveness detection methods that discriminate between human and machine sources for voice input require close speaking distance and thereby decrease usability, and analogous liveness strategies for other sensor and ownership data types remain an open area [13]. Finally, the survey concludes that future trends in mobile user authentication will move toward two or multi-factor schemes that “fuse” multiple metrics, including ownership factors, in an organic manner rather than simply appending tokens to passwords or biometrics. Inter- and intra-category comparisons across knowledge-, physiological-, behavioral-, and ownership-based techniques are used to argue that such fusion can improve security while preserving or even enhancing usability, provided that leakage of knowledge-based secrets and forgery of ownership credentials are addressed through more sophisticated integration and protection mechanisms [2,13].

2.3. Behavioral Biometrics as an Additional Dimension

2.3.1. Definition and Characteristics

Behavioral biometrics are grounded in measurable patterns of human actions rather than static bodily traits. Authentication mechanisms in this category rely on how users type,

move, speak, or interact with touchscreens and devices, and they are typically implemented within a pattern recognition framework. In mobile and ubiquitous settings, this class of biometrics exploits signals continuously available from embedded sensors and interaction logs, thereby supporting both identification and verification tasks without requiring explicit user attention at fixed time points.

A central distinction from physiological biometrics lies in the modalities and sensing infrastructure involved. Overviews of mobile authentication describe behavioral biometrics as encompassing keystroke dynamics, touch gestures, motion patterns, gait, voice, in-air handwriting, eye movement, and hand movement, sometimes further differentiated into subvariants such as text-dependent and text-independent keystroke or swipe gestures. These modalities become accessible on consumer devices through cameras, microphones, accelerometers, gyroscopes, and related sensors that collectively support both enrollment and ongoing verification. In this setting, behavioral signatures are captured while users perform ordinary actions, which can reduce intrusiveness compared with physiological traits that often require deliberate presentation.

Within behavioral biometrics, motion- and gait-based methods form a prominent subset. Surveys report modalities such as smartphone-based motion patterns, wrist or arm movements, and gait signals collected from accelerometers placed on the ankle, hand, hip, chest, or carried in the pocket [3,4]. Identity recognition schemes built around triaxial acceleration curves show that different activities, including walking, jogging, and sitting, generate distinguishable gait sequences suitable for classification. One proposed framework classifies gait types using a k Nearest Neighbor-based gait pattern classifier and then feeds the results into a convolutional neural network model that learns discriminative gait features before a softmax classifier outputs identity decisions [28]. These examples illustrate how behavioral biometrics often couple signal processing with machine learning to derive stable patterns from noisy, context-dependent motion data. Related work on motion-based identification from smartphone accelerometer streams shows that deep architectures combining ResNet and GRU backbones, together with Siamese and triplet-loss training regimes, can achieve competitive driver identification and verification performance on large, unconstrained datasets comprising dozens of drivers and tens of thousands of journeys, suggesting practical, lower-energy options for continuous authentication on resource-constrained mobile platforms [29].

Keystroke and touch dynamics represent another major behavioral category in mobile continuous authentication. Studies on keystroke-based systems transform raw keypress logs into timing features such as key-hold duration, inter-key intervals, and release-to-next-press delays, and then build per-user datasets that mix authorized and unauthorized events for supervised classification. Evaluations of mobile keystroke authentication compare multiple machine learning algorithms, including ensemble methods, k Nearest Neighbors, support vector machines, decision trees, and Naive Bayes, using metrics like accuracy and precision to quantify how well behavioral timing patterns separate legitimate from impostor events [30]. Broader behavioral surveys further catalogue motion-based, touch dynamics, and keystroke dynamics approaches, detailing conditions, classifiers, and performance metrics such as False Acceptance Rate, False Rejection Rate, and Equal Error Rate for continuous schemes [4,31].

From a system perspective, behavioral biometrics can be embedded in two main enrollment and decision paradigms. Template-based approaches collect several behavioral samples per user to construct reference templates, after which similarity measures against incoming samples determine acceptance or rejection. This strategy is often applied in physiological domains but can also be used for more stable behavioral traits. Model-based approaches instead train machine learning models directly from behavioral data, letting the

model decide whether new input belongs to the legitimate user. Common pipelines include data acquisition and preprocessing, feature extraction and selection, and classifier training, with feature quality playing a decisive role for performance. Evidence indicates that model-based methods have gained particular traction for behavioral user authentication due to their flexibility and ability to exploit complex sensor patterns [3].

Several cross-cutting characteristics emerge from these descriptions. Behavioral biometrics are typically less intrusive than many physiological modalities, since they leverage natural interaction rather than requiring explicit gestures for each authentication event. At the same time, they are inherently sensitive to context, device placement, and activity, which motivates multimodal combinations and adaptive schemes that can operate under diverse user circumstances. Continuous User Biometric Authentication Systems built on behavioral signals must therefore manage trade-offs between security and usability, taking into account False Acceptance and False Rejection behavior when designing thresholds and fusion strategies. Research on touch dynamics and related modalities also underscores the influence of usability, ergonomic, and human-based factors on recognition performance, prompting systematic evaluation of how interaction conditions shape behavioral signatures and their reliability in mobile environments [3,4,32]. Practical deployments have explored pairing behavioral signals with opportunistically collected device identifiers; for instance, a prototype using a Raspberry Pi to detect mobile device MAC addresses, capture vehicle plate images, and link these records to a backend and mobile application has been proposed to assist decision making in fuel stations [33]. Such integrated setups may suggest operational benefits for continuous authentication, automating correlation across interaction traces and device presence, but they also expose privacy, consent, and false-identification risks that require careful assessment before being folded into authentication pipelines.

2.3.2. Challenges in Behavioral Biometrics

Behavioral biometric approaches for continuous authentication encounter intertwined usability, security, privacy, and evaluation issues that complicate their deployment on mobile and ubiquitous platforms. Usability emerges as a central difficulty, since Continuous User Biometric Authentication Systems must control False Acceptance Rate and False Rejection Rate tightly enough to provide security without frustrating legitimate users through frequent denials. High FAR directly undermines protection, whereas elevated FRR degrades perceived usefulness, so system designers must explicitly balance these metrics when configuring behavioral schemes [4,9].

Security vulnerabilities span several behavioral modalities. Touch dynamics and keystroke dynamics have been shown to be vulnerable to generative attacks that deliberately manipulate machine learning models. Experimental evidence reports that adversarial generative attacks against touch dynamics can increase Equal Error Rate from 5% up to 50%, while similar attacks on keystroke dynamics can raise EER from 28% to 84%, severely degrading classifier reliability. Mimicry-style attacks constitute an additional threat: smartphone experiments demonstrate that keystroke dynamics cannot resist imitation attacks, and treadmill-based studies show that imitating user gait patterns can increase False Acceptance Rate from 5.8% to 43.66% in gait-based systems. Cardiac-based behavioral schemes are likewise at risk; if an attacker acquires a short electrocardiogram template, these data can be used to map the attacker's ECG into the victim's profile, with reported average attack success rates between 90% and 96%. Model-level weaknesses further exacerbate these threats. Many behavioral continuous authentication modes rely on machine learning classifiers such as Support Vector Machines and k Nearest Neighbors, often deployed in cloud-based settings. Investigative work shows that such models are exposed to sample inference, reconstruction, label distribution estimation, membership inference, model inversion, data

deanonymization, and model extraction attacks. For mobile-based continuous authentication in the cloud, reconstruction attacks have been demonstrated that identify users from stored positive samples in authentication profiles, indicating that models and stored data together form a rich target surface.

Privacy implications arise directly from the nature of behavioral signals. Gait recognition, stylometry, touch dynamics, and related modalities operate on daily routine data and current activities, so compromised profiles can reveal aspects of user behavior that, unlike passwords, cannot be permanently changed. Keystroke dynamics can expose demographic attributes such as age, gender, and the hand used for typing, while physiological modalities such as face recognition may disclose emotional states via facial expression analysis. Continuous monitoring and storage of such behavioral and physiological streams therefore threaten both identity privacy and contextual privacy when attacks succeed [9].

Continuous behavioral authentication also imposes significant resource and sensor related constraints on smartphones. Modern devices already struggle with battery life, and sensors such as light, proximity, gyroscope, and digital compass consume considerable power. Sensory data collection itself increases energy usage, yet CUBA operation requires sensors and processing pipelines to remain active throughout full sessions. Only a limited number of studies have examined smartphone power usage in detail, and conventional power management strategies that deactivate unused components conflict with the requirement for seamless behavioral monitoring, leaving energy efficiency as an unresolved challenge. One pragmatic response may be to push feature extraction and matching onto edge hardware and to adopt hardware-assisted countermeasures such as memory-centric and device-centric protections, which can cut network transfers and limit exposure of raw behavioral streams [34]. Yet these remedies are not free: they may shift complexity into device firmware and require careful evaluation of power, latency, and privacy trade-offs before they can be broadly recommended.

At the system design level, no single behavioral modality appears suitable for all user circumstances. The choice of modality must consider what the individual is doing or not doing at a particular time, yet many approaches still target narrow contexts. Keystroke dynamics is advocated because it is inexpensive, unobtrusive, and tolerant of different body positions, while mouse dynamics has been proposed to improve stability in desktop-style settings. However, most published work addresses one or two specific scenarios and seldom constructs a collection of modalities that can naturally select verification mode according to situation, which limits robustness across diverse real-world conditions.

Ethical and psychological aspects intersect with these technical issues. Continuous collection of sensitive behavioral data raises questions regarding consent, autonomy, and potential bias, and current systems rarely incorporate insights about users' emotional states or cognitive and social factors into their design. Suggestions for boosting usability include using psychological knowledge about behavior under different emotional conditions and conducting user surveys about experiences with various continuous methods, yet such interdisciplinary approaches remain under-explored in behavioral biometric research [4].

2.4. Behavioral Biometric Modalities

2.4.1. Keystroke Dynamics

As classified in the taxonomy of Figure 1, keystroke dynamics forms one of the four behavioral modalities examined in this review. Typing behavior on mobile devices provides a rich stream of temporal features.

Model construction for keystroke-based authentication typically follows a supervised learning pipeline. After feature extraction from keypress sequences, datasets are assembled that contain both authorized and unauthorized events, enabling training of binary or multi-

class classifiers. One comparative evaluation on mobile keystroke data considered ensemble techniques, *k* nearest neighbors, support vector machines, decision trees, and Naive Bayes, reporting performance in terms of accuracy and precision across these alternatives [30,35]. Within broader behavioral continuous authentication work, similar classifiers, including *k* Nearest Neighbors, Support Vector Machines, and Hidden Markov Models, are repeatedly employed across modalities such as keystroke dynamics, touch dynamics, and motion-based schemes [4,9]. This reliance on pattern recognition methods reflects the inherently statistical nature of keystroke features and their variability over time and context.

Continuous authentication scenarios impose additional constraints compared with static login checks. Rather than evaluating a single password entry, systems may monitor keystroke timing across prolonged sessions, potentially interleaving verification with other modalities when explicit typing is not available. SensorCA, for example, implements continuous verification of individual behavior patterns over sliding time windows by extracting sensor-based features and training a kernel ridge regression model with a tunable radial basis function, achieving an Equal Error Rate of 3.0% on datasets of size 1000 and 8000. Although SensorCA is not limited to keystrokes, it exemplifies how temporal behavioral streams can be segmented, classified, and evaluated in continuous mode using metrics such as False Acceptance Rate, False Rejection Rate, and Equal Error Rate [4,9]. Keystroke dynamics also appear in multimodal designs where wrist motion, mouse behavior, and typing patterns are fused to fill security gaps during short breaks in keyboard or mouse activity, highlighting their role as one component in richer behavioral profiles [4].

Security analysis exposes substantial vulnerabilities of keystroke-based continuous authentication. Generative adversarial attacks that manipulate classifier training or input data can drive Equal Error Rate from 28% up to 84%, demonstrating that learned decision boundaries can be severely compromised when exposed to crafted impostor samples. Mimicry attacks have also been demonstrated experimentally on smartphones, where adversaries attempt to imitate the typing style of a target; these studies conclude that keystroke dynamics cannot withstand such imitation, indicating that behavioral distinctiveness may not suffice against determined attackers. In gait-based systems, treadmill-assisted imitation can raise False Acceptance Rate from 5.8% to 43.66%, and analogous concerns about learnable behavior patterns apply to keystroke dynamics as well.

Machine learning infrastructure itself introduces further attack surfaces. Investigative work on continuous authentication models in cloud-assisted settings shows exposure to sample inference, reconstruction, and membership inference attacks, as well as model inversion, data deanonymization, and model extraction threats. When Support Vector Machines or *k* Nearest Neighbors are used for behavioral modes, actual user samples may be retained in authentication profiles; reconstruction attacks have successfully identified users from stored positive samples, illustrating how keystroke and other behavioral data, once uploaded, can be abused by adversaries with access to model internals or stored datasets. These findings indicate that keystroke-dynamics-based schemes must be evaluated not only for classification accuracy but also for resilience against attacks on both the feature space and the underlying learning algorithms. Indeed, empirical analyses appear to show that reconstruction attacks may remain feasible even when attackers lack direct access to original user samples, which suggests that cloud-assisted behavioral models should consider privacy-preserving mechanisms such as encrypted aggregation or federated updates [36]. This observation complicates simple reliance on accuracy metrics and highlights practical trade-offs between model utility and the need to limit exposure of raw timing traces.

Privacy implications accompany these issues. Behavioral biometric modalities authenticate by recognizing daily routine data and activity patterns, which cannot be reset like passwords once compromised. Keystroke dynamics in particular can leak sensitive demo-

graphic attributes, including user age, gender, and the hand used for typing; exposure of such characteristics through compromised profiles or inference attacks threatens identity and contextual privacy beyond mere account takeover [9]. Continuous collection and storage of keystroke timing therefore raise concerns about long-term traceability and profiling, especially in cloud-based continuous authentication where behavioral streams are centralized for model training and updating [4,9].

From a usability and operational perspective, keystroke dynamics offer several advantages but also important constraints. They are characterized as affordable to implement and non-obtrusive, with the keyboard identifier adaptable to different body positions, which supports deployment across varied device postures and form factors. However, keystroke-based schemes require sufficient typing activity to function; when users are not entering text, continuous verification must either fall back to other modalities or accept longer detection windows for intruders [4,37,38]. Furthermore, keystroke patterns are influenced by behavioral changes such as mood or stress, and keyboard layout shifts can disturb previously acquired models, implying that re-enrollment or adaptation strategies are necessary to maintain performance over time [4,22]. Research on behavioral biometrics in mobile contexts also notes that many published studies concentrate on narrow scenarios and do not provide modality collections capable of selecting verification modes dynamically according to current user activity, which limits the robustness of keystroke dynamics when deployed alone in diverse real-world settings [4].

Broader surveys of AI-enabled mobile security further note that deep learning methods have been applied to behavioral modalities such as keystroke verification to harden PIN and pattern entry against observation, while cautioning that such approaches introduce challenges in training time, scalability, privacy risk, and adversarial susceptibility that must be addressed in any deployed solution [39]. Evidence from large-scale keystroke research further indicates that recurrent deep models can learn free-text typing patterns across very large populations and achieve low equal-error rates on desktop keyboards while producing more modest results on mobile devices, with clear cross-device degradation that may imply device-specific retraining and enrollment requirements [40].

2.4.2. Touch Interaction and Swipe Patterns

Touch/swipe is the second of the four behavioral modalities positioned in the taxonomy of Figure 1. Touch-based interaction on smartphones yields a rich set of behavioral cues.

Tap and swipe gestures form a core subset of these touch-based modalities. Early work on touchscreen-based schemes extracted behavioral features from sliding operations, using x and y coordinates, touch pressure, touch size and time offsets to characterize user-specific swipe patterns. With Dynamic Time Warping (DTW) as the classifier, such a system achieved 77% accuracy with a 19% False Rejection Rate and 21% False Acceptance Rate on a dataset of 48 participants, illustrating that simple sliding gestures already contain discriminative information for user verification. Subsequent studies designed more complex multi-finger gestures, for example, static counter clockwise rotations with all five fingertips, and incorporated additional motion features from digital gloves, including angular values along x , y , and z axes together with roll, pitch and yaw. In one such system, combining three multitouch gestures yielded an Equal Error Rate of 1.58% on data from 34 users, while a single multi-touch gesture alone resulted in an average EER of 7.88%. Continuous authentication has also been realized purely from common navigational gestures such as horizontal and vertical strokes. One mechanism that exploits such touchscreen gestures reports an Equal Error Rate of 0% in intra-session tests, 2–3% in inter-session tests, and below 4% in authentication tests one week after enrollment, using k Nearest Neighbors

and Support Vector Machines with Gaussian Radial Basis Function kernels on data from 41 testers. Another scheme profiles keystroke, sliding, pinch and handwriting actions jointly, drawing on multiple touch related features including coordinates, pressure and size, and attains a minimum EER of 0.75% for sliding gestures, with all other action types remaining below 10% when classified using SVM with RBF kernel. These outcomes demonstrate that both simple and more elaborate swipe patterns can support highly accurate verification under realistic smartphone interaction conditions [41].

Survey work synthesizing these efforts positions touch gestures, including swipes, flicks, slides and handwriting, as a key behavioral modality for transparent authentication on a variety of touchscreen devices such as smartphones, smartwatches and digital cameras. Reported systems make extensive use of operating-system-level signals, extracting feature sets that encompass accelerometer readings, pressure, gravity, velocity, touch area and time related measurements. When such feature spaces are combined with distance-based classifiers, accuracy can reach 99% with minimal Equal Error Rates as low as 0.03%, highlighting the effectiveness of detailed swipe and touch statistics for user recognition. Studies further observe that augmenting touch gesture features with motion sensor data improves performance, since motion-aware swipe representations capture both on-screen contact and associated device micro-movements [2,3]. Work that aggregates multiple inertial datasets and trains deep networks indicates that phone carriage modes can be predicted with high accuracy, and a simple binary detector can first flag unfamiliar placements before finer-grained classification [42]. Incorporating this kind of location-aware preprocessing into touch-based authentication may therefore help mitigate session variability due to grip or placement, and could improve temporal stability and resistance to certain spoofing strategies.

From a broader behavioral biometrics perspective, touch dynamics are repeatedly examined alongside keystroke and signature modalities, with specific attention to how usability, ergonomic and human-based factors influence performance on mobile devices. One specialized survey compiles public datasets for swipe, keystroke and signature-based touch dynamics, documents evolving vulnerabilities and potential attack strategies against these systems, and discusses open research problems around the security usability trade-off and privacy concerns arising from continuous collection of detailed interaction traces. That analysis also catalogues fusion strategies for multimodality schemes and describes how factors such as hand posture, grip style and screen size impact the stability and distinctiveness of swipe and tap patterns, providing an empirical basis for designing swipe-based authenticators that remain both accurate and acceptable in everyday smartphone use [3,32]. Complementary work on gait recognition using smartphone inertial sensors has reported comparably strong performance: Siamese neural network variants attained Rank-1 identification rates of about 96% and Rank-5 near 98%, while IMU-driven deep convolutional approaches have reported identification accuracies on the order of 90–92% [43]. Such results may suggest that judiciously fusing touch gesture statistics with inertial-sensor gait cues could improve the temporal stability and attack resilience of continuous authentication systems, although session variability and sensor placement effects are likely to remain practical challenges.

2.4.3. Physical Activity and Motion Sensors (Gait)

Gait/motion is the third of the four behavioral modalities positioned in the taxonomy of Figure 1. Gait is defined as the walking posture of a subject.

Two main methodological families can be distinguished for sensor-based gait recognition. Template matching approaches construct gait templates under normal conditions and then recognize a user by computing similarity between stored templates and incoming gait cycles, using measures such as Dynamic Time Warping, Pearson Correlation Coefficient or

Cross Correlation. These methods have achieved good performance in laboratory settings, but require accurate gait cycle detection to build both reference templates and test samples. Detecting gait cycles is challenging because it is sensitive to noise, device position, pace, and road conditions, and failures or inter-cycle phase misalignment can lead to incorrect recognition decisions. Machine learning-based methods instead extract features from inertial signals and classify them into user identities using models such as Support Vector Machines, k Nearest Neighbors and Random Forests, generally achieving more robust performance and reducing dependence on handcrafted similarity measures [44,45].

Deep learning has become a prominent trend in IMU-based gait recognition, motivated by the limitations of manual feature engineering. Studies using Convolutional Neural Networks within frameworks like IDNet demonstrate that CNNs can automatically learn useful features from smartphone accelerometer and gyroscope data, outperforming manual feature extraction combined with classical classifiers [44]. Subsequent work proposes a lightweight attention-based CNN that employs contextual encoding and depthwise separable convolutions to reduce model parameters while strengthening extracted gait features, targeting deployment on wearable devices with constrained computational and memory resources [3,44]. Experiments on the large-scale OU ISIR dataset with 744 participants and the whuGait dataset, which contains inertial data from 118 subjects collected in unconstrained real-world conditions, show that such lightweight models can improve recognition accuracy compared with existing schemes while remaining suitable for wearable implementation [37,44].

Continuous authentication scenarios further motivate gait-based schemes that operate transparently across multiple devices. Motion sensor data from accelerometers and gyroscopes embedded in smartphones, smartwatches, tablets and smart bracelets can capture movement and posture, and bibehavioral features derived from these signals are identified as promising for multi-device continuous authentication. Tablet-based device surveys document that accelerometers are available on all major smart devices, while gyroscopes are present on most of them, and that these devices can communicate with cloud servers or mobile terminals via mobile networks, WiFi or Bluetooth, enabling distributed gait authentication architectures. Research on cross-device continuous authentication has explored correlations between keystrokes and touchscreen biometrics across smartphones, tablets and desktops, but extending continuous schemes to smart wearables requires motion sensor-based features such as gait that are not tied to keyboards or touchscreens [46].

Within mobile security, gait recognition is typically realized using wearable or embedded motion sensors to support continuous and transparent authentication as the user walks, following a pipeline of data acquisition, preprocessing, walk detection and analysis. Acquisition strategies often constrain device placement, for example, in a pouch, trouser pocket or hand, to obtain consistent motion readings, though this can reduce ecological realism [3]. Studies survey that mobile gait datasets commonly involve population sizes of 50 participants or fewer and are often collected under controlled conditions, which raises questions about scalability and robustness in broader deployments [3,44]. In contrast, one large-scale study on gait recognition using smartphones in unconstrained real-world conditions reports strong identification performance across a significantly larger and more diverse population, demonstrating that deep networks trained on wild data can generalize beyond controlled laboratory settings [47]. Complementary work on continuous authentication using gait biometrics reviews schemes that employ triaxial accelerometers worn at the wrist or thigh, apply autocorrelation-based filters to discard unsuitable segments, and then use anomaly detection to separate genuine users from impostors, achieving Equal Error Rates as low as 2.5% in experiments with 20 volunteers [4].

Broader analyses of continuous authentication techniques that utilize gait highlight both advances and unresolved issues. Reported methods exploit unique patterns in an individual's walking style for security and authentication, and some achieve EER values below 1% in same-session evaluations, yet long-term temporal robustness and performance under varied environmental conditions remain insufficiently studied [4,44]. Challenges identified for motion-based schemes include optimizing computational and memory requirements on mobile and wearable platforms, and addressing adversarial attacks such as observation-based imitation and sensor-based side channel inference, which can threaten gait-based authentication despite its continuous and ostensibly hard to forge nature [3,4]. Research on ambient smart applications further demonstrates that wearable accelerometer-based gait signals can authenticate users passively across everyday environments, achieving stable recognition performance without requiring dedicated enrollment sessions, which reinforces the suitability of gait biometrics for continuous, unobtrusive mobile authentication [48].

2.4.4. Contextual Data

Contextual data completes the four-modality decomposition of the behavioral-biometrics branch shown in Figure 1. Contextual information augments behavioral biometrics.

Location in particular appears as a recurring contextual factor when combined with behavioral biometrics. Continuous authentication frameworks for mobile applications integrate user location data alongside movement information and device usage patterns, often constructing separate profiles for weekdays and weekends to capture routine mobility differences. One multimodal proposal clusters GPS locations using the K means algorithm, yielding per-user spatial patterns that complement other behavioral cues. Another architecture blends GPS information with stylometry, application usage statistics, and web search patterns, thereby embedding location into a richer behavioral context. In both cases, location is treated as relatively stable for many application use scenarios, and deviations from established location patterns contribute to detecting anomalous access in mobile continuous authentication. Contextual features are thus not limited to instantaneous coordinates but can encode long-term usage habits that strengthen behavioral profiling [49].

Research on continuous authentication and user profiling frames these contextual modes within a wider taxonomy of biometric and non-biometric modalities. A comprehensive survey distinguishes context-based, behavioral, and physiological biometrics and explicitly includes context-aware techniques as one of the core categories analyzed. Within this structure, context-aware authentication is defined as dynamically adapting authentication mechanisms based on environmental factors and user behavior patterns, such as location, time of day, and activity. By integrating such contextual information, systems can vary authentication strictness and trigger additional checks when deviations from established patterns are observed, thereby aiming for a more seamless and secure user experience. The same review compiles a dedicated table of context-aware continuous authentication works, organizing them by source, classifier, performance metrics, and specific contextual features used, which demonstrates that context signals are being systematically evaluated alongside more traditional biometric traits [4].

Despite these advantages, significant challenges accompany the use of contextual data in continuous authentication. Security subsections of the survey note that context-aware modes are exposed to vulnerabilities, particularly regarding data privacy and unauthorized access. For example, systems that depend on GPS information may not always obtain accurate or trustworthy location readings, which can lead to misclassifications and open avenues for spoofing through manipulated context inputs. Broader discussions of security and privacy emphasize that continuous monitoring of contextual attributes, along with biometric signals, may infringe on individual privacy rights if deployed without transparent

policies and robust protection measures. Users might be unaware of how extensively their contextual information is collected and used for authentication, which motivates explicit consent management and clear disclosure of data handling practices. Ethical analyses therefore stress that developers must treat contextual data with the same care as biometric data, addressing concerns about consent, data usage, and potential bias when embedding context-aware modes into continuous authentication frameworks [4].

Energy consumption imposes a further constraint on context-dependent systems operating on smartphones. Many contextual features rely on sensors such as GPS, light, proximity, gyroscope, barometer, accelerometer, and digital compass, as well as continuous logging of network and application usage. Evidence indicates that these sensors consume substantial battery power and that sensory data collection itself incurs high energy costs. Only a limited number of studies provide detailed analysis of smartphone power usage, yet general power management strategies that deactivate sensors when idle conflict with continuous authentication requirements, where sensors and processing must remain active for the entire session. This tension is particularly pronounced for context-aware behavioral schemes that require always-on monitoring of environmental and usage factors to maintain security, underscoring the need for more energy-efficient designs and careful modality selection in resource-constrained mobile settings [4,9].

Synthesis across modalities. Read together, the four modalities reviewed above show a recurring conflict rather than a consistent picture: reported error rates are lowest precisely where sample sizes and environmental realism are smallest. Same-session gait evaluations with 20 volunteers report EER as low as 2.5%, and some same-session designs report sub-1% EER [4,44], yet the one large-scale, unconstrained-condition gait study surveyed here [47] evaluates a substantially larger and more diverse population and does not report comparably low error rates under real-world conditions, and neither controlled small-cohort study demonstrates that its reported error rate is stable outside the laboratory. This same small-cohort pattern recurs in keystroke, touch, and contextual studies above, most of which draw on cohorts of well under 50 participants collected in short, controlled sessions rather than longitudinal, ecologically valid deployments. A second recurring limitation is energy: every modality surveyed here, from motion sensing to always-on contextual logging, is reported as power-intensive, yet almost none of the underlying studies measure energy consumption directly, instead treating it as an acknowledged but unquantified constraint. A common methodological trend follows from both observations: studies increasingly fuse modalities (touch with motion, GPS with app-usage statistics) specifically to compensate for the fragility of any single modality's small-cohort performance, rather than because fusion was independently motivated by a fusion-specific hypothesis. This suggests that single-modality error rates reported in isolation in Table 3 should be read as upper bounds on realistic performance, not as deployment-ready benchmarks.

Table 3. Behavioral biometric studies with reported performance metrics.

Reference	Modality	ML Method	Dataset	Performance	Key Contribution
[40]	Keystroke	LSTM (TypeNet)	Aalto Mobile Keystroke	EER: 9.2%	Introduces TypeNet for free-text keystroke biometrics using deep learning.
[23]	Keystroke	Transformer	Aalto Mobile Keystroke	EER: 3.84%	Introduces transformer-based mobile keystroke biometrics for continuous authentication.
[50]	Keystroke	Classical ML	Proprietary	EER: 0.72%	Demonstrates high-security authentication using piezoelectric keystroke dynamics.
[30]	Keystroke	Classical ML	Proprietary	Acc: 85%	Compares ML classifiers for continuous authentication on mobile devices by keystroke dynamics.

Table 3. Cont.

Reference	Modality	ML Method	Dataset	Performance	Key Contribution
[51]	Touch	Classical ML	Proprietary	Acc: 98%	Proposes swipe-based unlocking with supervised learning on smartphones.
[45]	Gait	CNN	Proprietary	Acc: 92.43%	Evaluates deep learning approaches for continuous authentication via activity patterns.
[43]	Gait	Siamese NN	Proprietary	Acc: 96% (Rank-1)	Evaluates Siamese neural networks for gait recognition with wearable sensors.
[52]	Gait	CNN-Bi-LSTM	Proprietary	Acc: 98.4%	Proposes hybrid CNN-Bi-LSTM for continuous mobile user authentication.
[37]	Multimodal	Deep Learning	Proprietary	TAR: 99.98%	Reviews deep learning and ML together for biometrics mobile authentication.
[53]	Multimodal	Deep Learning	Proprietary	TPR: 99%	Implements DAKOTA continuous authentication on a mobile banking application.
[54]	Multimodal	RNN	Proprietary	EER: 3.8–11%	Applications of RNN for biometric authentication & anomaly detection.
[46]	Multimodal	Deep Learning	Proprietary	Acc: 99.8%	Proposes continuous user authentication across multiple smart devices.
[47]	Gait	Deep Learning	OU-ISIR/whuGait	Acc: 90%+	Deep learning-based gait recognition using smartphones in the wild.
[55]	Keystroke	Deep Learning	Proprietary	EER: <1%	Proposes freely typed keystroke dynamics using heterogeneous features.
[56]	Face	CNN	NUAA/CASIA	Acc: 94.17%	Proposes multi-layer CNN for facial spoofing detection and classification.
[57]	Face/Iris	CNN	LivDet-Iris 2020	EER: 0.16%	Presents MobileNetV2 architecture for iris liveness detection.

Note: EER—Equal Error Rate; Acc—Accuracy. Color coding: Keystroke , Touch , Gait , Multimodal , Face . Only studies with reported performance metrics are included (n = 16).

2.5. Intelligent Personal Agents in Authentication Systems

2.5.1. Role of the Personal Agent

Continuous authentication research positions intelligent, agent-like components as decision makers that must reconcile competing requirements of security, usability, privacy, and ethical acceptability. Survey work on Continuous User Biometric Authentication Systems argues that usability trade-offs, particularly in terms of False Acceptance Rate and False Rejection Rate, are central to acceptance and must be considered at every development stage. Within such systems, a personal agent that coordinates sensing, feature extraction, and decision thresholds effectively becomes responsible for implementing these trade-offs in real time, since it must determine when stricter checks are justified and when user burden should be minimized to avoid frustration and reduced usability.

Security responsibilities of an intelligent personal agent extend across multiple biometric and contextual modalities. Comprehensive analysis of continuous authentication reports that behavioral and physiological biometrics, as well as context-aware techniques, remain vulnerable to adversarial attacks and spoofing, including generative attacks against touch dynamics and keystroke dynamics that dramatically increase Equal Error Rates. A personal agent orchestrating continuous authentication must therefore monitor for such degradation in classifier performance and, where possible, adapt model usage or trigger fallback strategies when evidence of spoofing or adversarial manipulation arises. At the same time, context-aware modes, for example, those using GPS information, introduce specific risks when contextual data are inaccurate or manipulated, which again requires

agent-side logic capable of handling uncertain or inconsistent inputs without compromising security [4,9].

Privacy and ethical stewardship form another core dimension of the personal agent's role. Continuous operation implies persistent collection and storage of biometric and contextual data, and analyses warn that such monitoring can infringe individual privacy rights, particularly when users are unaware of the extent of data collection and subsequent usage. Ethical discussions stress that any system relying on sensitive biometric and behavioral data must address consent, data usage, and potential bias, with developers adhering to principles that prioritize user autonomy, privacy, and fairness. In practice, this places the personal agent at the center of enforcing privacy-preserving mechanisms, for instance, by restricting data retention or controlling access to sensitive templates, and of mediating transparent communication with users about how authentication data are processed [4].

Several studies point to concrete technical responsibilities that such agents must assume in mobile and cloud-assisted environments. Work comparing cloud-side and device-side continuous authentication highlights that mechanisms may operate locally on the device or involve transmission of biometric and context data to remote servers for processing, with each choice creating distinct privacy and security challenges. When processing is offloaded, a curious insider at the server side may obtain access to behavioral profiles, templates, and contextual streams, so the personal agent must manage data flows and decide which information remains on device and which can safely be transmitted [9].

Surveyed privacy-preserving approaches, including zero-knowledge-proof-based protocols and other cryptographic techniques, are explicitly proposed to reduce what servers can learn about user secrets, which again requires an agent capable of invoking such protocols and ensuring that authentication succeeds without disclosing keys or explicit authentication information to remote entities [4]. Research on agent workflow systems highlights an absence of standardized specifications and suggests protocols such as Agent2Agent to enable clearer role definitions and interoperable interfaces for coordinating tools and data exchanges, which may be instructive when designing how personal agents select modalities, negotiate processing boundaries, and record provenance [58]. The same literature also flags internal threats like model contamination, tool poisoning, and memory-based privacy leaks, implying that continuous-authentication agents should integrate monitoring, input sanitization, and memory hygiene measures rather than relying solely on access controls or encryption.

Memory-based leakage deserves particular attention for continuous-authentication agents specifically, since these agents necessarily retain behavioral templates, recent authentication decisions, and contextual state across a session in order to perform sequential, multi-step verification rather than a single point-in-time check. This session-level memory is precisely the kind of internal state that broader agent security surveys flag as vulnerable: information retained in an agent's working memory or context can be exposed through improper memory management, shared caches, or interaction with external tools and sub-agents that the authentication agent invokes, even when no single component is individually compromised [34]. For an authentication agent specifically, this implies that template confidentiality cannot rely solely on the TEE boundary introduced in Section 2.5.2: any behavioral data or decision state that the agent passes outside the TEE for logging, debugging, or coordination with edge or cloud components becomes a candidate memory leakage point, regardless of how well the core model and decision logic are isolated. Concrete mitigations discussed in the broader agent security literature, including scoped, time-limited retention of session state and explicit clearing of behavioral data once a decision is finalized, are directly applicable here but have not yet been evaluated specifically

for continuous behavioral authentication agents, leaving this as an open methodological gap rather than a solved problem.

Translating the three-tier architecture in Figure 2 into an actual deployment raises several practical considerations that are distinct from, though related to, the modality-level performance and security analyses presented elsewhere in this review. First, the TEE-based device tier assumes hardware that not all mobile devices provide: while flagship smartphones commonly include a dedicated secure enclave or equivalent, lower-cost and older devices may lack one, which means a deployable system must define a fallback behavior, such as reduced-confidence operation or mandatory escalation to edge-tier verification, for devices without TEE support rather than assuming uniform hardware capability. Second, mobile operating systems increasingly restrict background sensor access and continuous processing to conserve battery, through mechanisms such as app standby and power-saving modes; a continuous-authentication agent that requires uninterrupted sensor streams (Section 2.1) must therefore negotiate with, rather than bypass, these OS-level energy management policies, which in practice constrains how literally “continuous” the monitoring can be on a stock consumer device. Third, the model update flow shown as a single arrow between the cloud and edge tiers in Figure 2 corresponds in practice to an ongoing maintenance pipeline: behavioral drift (RQ4, Section 1.1) means models cannot be deployed once and left static, so a real deployment must budget for periodic retraining, versioning, and rollback procedures, with the same privacy-preserving mechanisms discussed for initial template protection (cancelable biometrics, federated learning) extended to cover the update process itself. None of these constraints invalidate the architecture, but they indicate that moving from the synthesized framework to a production system requires explicit engineering decisions that the modality- and security-focused literature reviewed in this paper does not yet address in a unified way.

As a coordinator of heterogeneous modalities, the personal agent also has to manage modality-specific limitations and context-dependent suitability. Evidence shows that no single biometric modality is appropriate for all user circumstances, and that the individual’s activity at a given time should guide modality selection. Keystroke dynamics is described as affordable, unobtrusive, and tolerant of different body positions, whereas mouse dynamics is suggested to enhance stability in some settings, yet both modalities remain insufficient when the user is not typing or interacting with a pointer device. Research explains that future directions should focus on advancing biometric technologies, integrating multimodal approaches, and embedding context-aware techniques so that systems can adapt to diverse user scenarios and leverage machine learning and artificial intelligence to improve accuracy and reliability. Within this vision, a personal agent must act as the orchestration layer that selects appropriate modalities based on current context, fuses their outputs, and exploits advanced learning methods to maintain robust authentication while respecting the constraints of each sensor and activity state.

Resource management is another area where personal agents assume an operational role. Smartphones rely on multiple sensors, including light, proximity, gyroscope, and digital compass, and continuous data collection requires substantial energy, with only a few studies providing thorough investigation of smartphone power usage [4,9,59]. Typical power management would turn off unused components, yet continuous authentication demands seamless processing and monitoring with sensors powered throughout active sessions. An intelligent agent embedded in such a system must therefore balance continuous verification requirements against battery constraints, potentially prioritizing lower cost modalities or adjusting sampling strategies to preserve energy without sacrificing essential security guarantees [4,9].

Finally, agent-like components are central to future research directions that seek to integrate emerging biometric modalities, advanced machine learning, and novel privacy-preserving technologies. Suggested avenues include exploring new traits such as motion-based biometrics, adopting multimodal and context-aware authentication strategies, and incorporating techniques like zero-knowledge proof and quantum cryptography to safeguard privacy [4]. Research on behavioral biometric continuous authentication additionally highlights that user perceptions, including perceived vulnerability, severity of consequences, self-efficacy, and response costs like power consumption, shape intentions to adopt such technologies [1]. In this landscape, the role of the personal agent includes not only technical decision making but also mediating between system capabilities and user attitudes, aiming to deliver continuous authentication that users judge both effective and acceptable in everyday mobile use [1,4].

2.5.2. Architecture of the Intelligent Agent

Architectural designs for intelligent agents in authentication and security contexts typically couple sensing, decision making, and policy enforcement within a distributed infrastructure that spans constrained devices, edge nodes, and cloud services. In trust-based monitoring for cloud-assisted IoT communications, middleware and intelligent agents collaborate in three phases, namely spoof detection, trust construction, and message authentication, with the agents exchanging trust and signal strength observations and handling monitoring, processing, and task switching to reduce communication cost and energy consumption [7]. Similar multi-tier arrangements appear in blockchain enabled smart home access control, where end users, IoT devices, multi-edge servers, and cloud servers all interact with Ethereum smart contracts using unique public and private keys. These designs illustrate that an intelligent agent responsible for authentication rarely resides in isolation on a single device; instead, it is embedded in a layered architecture that must balance local responsiveness with offloaded computation and long-term data management. Within smart home scenarios, edge nodes act as administrative agents that interface directly with sensors and actuators, maintain local storage, and manage smart contracts that register users and devices. Edge servers use a shared key for local communications, process all incoming and outgoing transactions, authenticate end users to access IoT devices, and perform mining operations on behalf of resource-constrained nodes, propagating data to the cloud for further analytics under a differential privacy enhancement mechanism [60]. This arrangement effectively partitions responsibilities: low-level sensing and actuation occur at IoT devices, policy evaluation and transaction validation sit at multi-edge agents, and large-scale data analysis and archival storage are delegated to cloud components. Trust-based monitoring in IoT adopts a related pattern in which intelligent agents reside near devices, exchange attribute and signal observations with middleware, and collectively maintain security state that informs message authentication and spoof detection decisions [7]. Together, these examples define a reference architecture in which the intelligent authentication agent spans at least a device or gateway tier and an upper aggregation tier. Figure 2 illustrates this three-tier design, depicting how sensing, local processing, edge-level security enforcement, and cloud-based model training and privacy preservation are distributed across device, edge, and cloud layers.

Within the device tier shown in Figure 2, the interaction between the agent, the sensors, and the TEE follows a specific sequence. Embedded sensors (accelerometer, gyroscope, touch screen, keyboard) continuously stream raw behavioral signals to the local agent's feature extraction stage, which runs in the normal, untrusted execution environment of the operating system since it primarily performs signal processing rather than security-critical decisions. The resulting feature vectors are then passed into the TEE boundary, where the

local ML model and the authentication decision logic execute in an isolated, hardware-backed enclave that the host operating system, and any malware running on it, cannot directly inspect or modify. This separation means that even if the device's main operating system is compromised, an attacker cannot read out the behavioral templates held by the local model or tamper with the logic that produces accept or reject decisions, though the feature extraction stage upstream of the TEE remains exposed and could in principle be manipulated to feed corrupted inputs into the enclave. The authentication decision produced inside the TEE is what is ultimately encrypted and forwarded to the edge tier for further processing, rather than raw sensor data or unprotected model parameters.

Continuous authentication adds further requirements to this agent architecture. Case analyses of continuous authentication describe three main processing configurations: computation performed in the cloud for cloud service access, computation executed on the device for device access, and computation offloaded to the cloud for authenticating users to the device. Each case implies a distinct placement of the core decision logic of the intelligent agent and creates different trade-offs between power consumption, communication overhead, and exposure of personal data. When processing resides in the cloud, the agent must orchestrate continuous data transmission from devices and tolerate high communication cost, while facing privacy risks from revealing behavioral and contextual data to remote servers. When computation remains on device, the agent must manage substantial memory and compute requirements and cope with battery constraints, since CUBA operation requires that sensors and processing pipelines remain active throughout active sessions [4,9]. In hybrid arrangements, a local component handles data acquisition and preliminary processing, whereas remote services perform heavier classification, but privacy concerns reappear due to outsourcing.

Threat actor assumptions strongly influence architectural responsibilities of intelligent agents. For cloud-based continuous authentication, analyses posit a curious insider on the server side who can access location data, IP addresses, online activities, and even partial or full templates, feature computation details, and selection criteria, and who can reconstruct biometric templates given sufficient information [9]. Under such threats, an intelligent agent that orchestrates authentication must include mechanisms for limiting what is transmitted, for applying privacy-preserving transformations, and for managing template protection across storage locations. Survey work on continuous authentication highlights the need to integrate techniques such as zero-knowledge proof protocols and other cryptographic schemes so that servers cannot learn cryptographic keys, metadata, or explicit authentication information, even when they participate in verification [4]. In healthcare-oriented IoMT settings, an eHealth authentication framework with sensors, medical experts, a mobile sink, and an authentic gateway assumes that the gateway acts as a trusted node managing authorization, while key generation hardware and secret key distributions enforce mutual authenticity and single sign on across real-time entities [8]. Here, the gateway effectively embodies the central intelligent agent, mediating between distributed sensors and semi-trusted cloud servers that may collude with revoked users.

Blockchain-based architectures provide another structural option for intelligent agents in authentication ecosystems. In smart homes, attribute-based access control is realized via Ethereum smart contracts deployed on multi-edge servers, with each participant holding a distinct address and key pair [60]. Edge agents handle contract maintenance, user and device registration, and local authentication decisions, while also performing mining to support the blockchain ledger. In vehicular and IoV contexts, blockchain-assisted cross-domain authentication uses consensus algorithms such as PBFT and smart contracts to manage pseudonym issuance, revocation, handover authentication, and roaming, storing auxiliary authentication data on-chain to improve certification efficiency and resist distributed de-

nial of service attacks [61]. Cross-domain anonymous authentication models depict base stations, trusted authorities, and consensus servers interconnected through blockchain, with fog or edge nodes acting as miners and authentication relays, again distributing the intelligence of the agent across several coordinated components.

Energy and performance constraints on mobile and IoT devices further shape how an intelligent authentication agent must be architected. Smartphones rely on multiple sensors, including light, proximity, gyroscope, and digital compass, and sensory data collection increases power consumption significantly; only a few studies provide thorough analysis of smartphone energy usage, yet typical power management strategies that deactivate unused components conflict with continuous monitoring requirements [4]. Surveys of mobile biometric modalities indicate that physiological traits such as fingerprint, iris and face generally yield lower false rejection rates and have been widely adopted on devices, while behavioral and physiological signals like swipe dynamics, ECG and EEG can be collected more passively but often carry higher variance in accuracy [56]. Architecturally, this implies an agent must be aware of modality-specific acquisition and preprocessing needs, temporal feature extraction for touch, resolution and illumination constraints for iris, or cycle-count requirements for ECG, so it can schedule sensing, local filtering, and selective transmission to save energy without eroding authentication utility. In trust-based monitoring for IoT, intelligent agents explicitly assist with monitoring, processing, and task switching to reduce communication costs, achieving lower response and detection times, lower misdetection probabilities, and reduced false positive rates while improving network lifetime through decreased energy consumption [7]. Federated learning frameworks for healthcare similarly distribute learning across devices, allowing local models to be trained at sensing nodes and only weights and gradients to be shared with centralized servers, thereby reducing computation complexity at the center and preserving device privacy, although untrusted servers may still exploit model parameters to recover private data [8]. In such environments, an intelligent agent must coordinate local model training, secure aggregation, and pseudonym updates, while respecting constraints on computation, communication, and battery.

Finally, architecture level considerations must incorporate usability, security, privacy, and ethical requirements already identified for continuous authentication systems. Comprehensive reviews stress that CUBASs must tune thresholds and operating points to balance False Acceptance and False Rejection Rates, ensure that behavioral and physiological modalities are protected against spoofing and adversarial attacks, and handle privacy implications from continuous collection of biometric and contextual data [4,9]. Intelligent agents situated within these systems therefore not only execute classification and access control logic but also embody policy decisions about data retention, consent handling, and modality selection in response to user activities. Research directions that advocate integrating multimodal biometrics, context-aware techniques, and privacy-preserving protocols such as zero-knowledge proof and blockchain-based transparency mechanisms implicitly assume an architectural layer where such diverse components are orchestrated coherently [4,60,61].

Synthesis across agent architectures. The architectures surveyed above disagree less on which components are needed (sensing, local feature extraction, a decision layer, and a privacy-preserving transport mechanism) than on *where* the trust boundary is drawn: cloud-centric designs place trust in the server operator and accept a curious-insider threat model [9], on-device and TEE-based designs place trust in device hardware and accept exposure of the feature-extraction stage upstream of the enclave, and blockchain-mediated designs place trust in consensus rather than in any single party but introduce mining and consensus overhead as a new resource cost. No single architecture reviewed here

resolves all three constraints (privacy from the server, integrity on the device, and low energy/communication cost) simultaneously; each design surveyed trades one for the others. A second recurring limitation, common to the cloud, edge/TEE, blockchain, and federated designs alike, is that energy and communication overhead is repeatedly flagged as a constraint but is rarely measured empirically for the specific agent architecture proposed, echoing the same measurement gap identified for individual behavioral modalities above. The clearest methodological trend is a shift from single-tier (all-cloud or all-device) designs toward explicitly hybrid, multi-tier orchestration, suggesting that the field is converging on tiered architectures as a practical compromise, even though, as noted above, no reviewed study evaluates such a hybrid pipeline end-to-end under simultaneous adversarial and privacy threat models.

2.6. Machine Learning Methods

2.6.1. Classical Algorithms

Pattern recognition in continuous and mobile authentication has long relied on established machine learning algorithms such as Support Vector Machines, k Nearest Neighbors, Random Forests, kernel ridge regression, and various neural network variants. These methods underpin many behavioral biometric and multimodal schemes that predate or complement recent deep learning approaches. In biometric gait recognition, for example, template matching with Dynamic Time Warping and correlation-based distances has been widely used, while classical classifiers such as kNN, SVM, and Random Forests have provided more robust performance when built on manually engineered features from inertial sensors [3,44].

Keystroke dynamics and touch-based authentication on mobile devices likewise depend heavily on classical supervised learning pipelines. Raw keypress logs or swipe traces are transformed into timing and geometric features, after which algorithms including kNN, SVM, decision trees, Naive Bayes, ensemble methods, and kernel ridge regression are trained on labelled authorized and unauthorized events. Comparative evaluations on mobile keystroke datasets report accuracy and precision for such classifiers and use metrics like Equal Error Rate to quantify their suitability for continuous verification. Similar strategies appear in swipe and pattern lock studies, where one class SVM, Gaussian mixture models, Random Forests, and other traditional learners are benchmarked across multiple public datasets and usage scenarios, often revealing that classifier choice interacts strongly with user stability and gesture type [3,30,37,62].

Classical algorithms are also central in multimodal and context-aware continuous authentication. SensorCA, a representative continuous scheme, uses kernel ridge regression with a radial basis function kernel on features derived from multiple hard and soft sensors, achieving an Equal Error Rate around 3% on datasets of size 1000 and 8000 when operating over sliding time windows. Other multimodal frameworks fuse outputs from SVM, Random Forest, and Naive Bayes classifiers operating on touch dynamics, motion signals, WiFi usage, and application logs, in some cases reaching authentication accuracies above 99% but at the cost of increased computation and memory overhead on the device [3,9,63].

Performance comparisons consistently document strengths and limitations of these classical learners. In gait recognition using wearable IMU sensors, Random Forests and SVMs outperform template matching approaches but still depend on carefully hand-crafted features and extensive feature engineering, which is described as time consuming and reliant on domain expertise. Similarly, in touch stroke authentication, SVMs often provide the lowest average error rates compared with decision trees, back propagation neural networks, and Naive Bayes, leading some designs to adopt SVM as the default classifier in user studies once its advantage is empirically established. For pattern-based unlocking

mechanisms, one class SVM frequently surpasses k means and other clustering baselines, particularly when context awareness and fine-grained feature segmentation are applied to motion enhanced pattern traces [37,44,51].

Alongside these accuracy gains, several challenges arise when deploying classical algorithms in continuous, mobile settings. Many behavioral schemes run on smartphones equipped with accelerometers, gyroscopes, magnetometers, light, and gravity sensors, and continuous feature extraction plus classification at high sampling rates can impose significant computation and energy costs. Studies of multimodal continuous authentication explicitly note that running multiple models in parallel for different modalities introduces resource bottlenecks, even though modern phones now provide multi-core CPUs, GPUs, and gigabytes of RAM. Moreover, classical models such as SVM and kNN often store actual user samples in authentication profiles, which investigative work shows can be abused through reconstruction attacks in cloud-based deployments; positive samples belonging to a single user have been leveraged to re identify individuals, exposing both security and privacy risks that go beyond mere classification errors [3,9].

Security analyses further demonstrate that classical classifiers used in behavioral biometrics are vulnerable to adversarial and generative attacks. Touch dynamics systems that employ traditional machine learning can suffer Equal Error Rate increases from 5% up to 50% when exposed to crafted adversarial samples, while keystroke dynamics schemes show EER degradation from 28% to 84% under similar attack models. These results underline that even when FAR and FRR are acceptable under benign conditions, classical algorithms may fail catastrophically when adversaries manipulate training data or input distributions. Additional work documents mimicry attacks against keystroke-based schemes and treadmill-assisted imitation attacks on gait authentication, both targeting the behavioral patterns on which SVM, kNN, and related classifiers are trained; in gait systems, such imitation can inflate False Acceptance Rate from 5.8% to 43.66% [4,9].

Despite these vulnerabilities, classical machine learning remains integral to continuous authentication research synthesized in Section 2.3.1. Surveys on continuous authentication and user profiling explicitly structure their analyses around physiological, behavioral, multimodal, and context-aware methods, and repeatedly report SVM, kNN, Random Forests, kernel ridge regression, Hidden Markov Models, and simple neural networks as the dominant algorithmic choices. These methods provide a well understood baseline against which newer deep architectures are compared, and they continue to be attractive for resource-constrained devices and scenarios where labelled data are limited or feature spaces are modest in size [3,4,9,37]. A recent study of freely typed mobile keystroke dynamics highlights how fusing heterogeneous low-level signals (accelerometer, touch coordinate and timing) can drive classical systems to unexpectedly low error rates: the FACT model reports EERs below 1% with a single reference set and near-zero EER with multiple reference sets, and the authors have released their dataset for further evaluation [55]. Such findings suggest that, in some cases, careful feature engineering and sensor fusion allow traditional learners to approach practical usability without resorting to deep models, although real-world variability and attack resistance remain open concerns. For instance, experiments on a piezoelectric touch dataset report Random Forest and neural network systems achieving EERs down to 0.72% while incurring only a few milliseconds per test decision on commodity CPUs, which appears to support the view that lightweight classical models combined with well chosen features may suffice for many mobile keystroke tasks [50].

2.6.2. Deep Learning

Deep learning is increasingly applied to biometric authentication because it can learn hierarchical representations from large amounts of data, including unlabeled datasets,

which reduces the need for manual feature engineering. In continuous and mobile authentication, convolutional neural networks, recurrent neural networks, and deep neural networks have been used alongside traditional machine learning algorithms to process both physiological and behavioral biometric signals. These architectures are particularly attractive where time-series or sequential interaction data must be modeled continuously during device use, such as touch, motion, and keystroke streams.

For behavioral keystroke dynamics, several works employ RNN-based deep architectures, most commonly long short term memory networks. These findings show how deep temporal models benefit from multimodal fusion in mobile continuous authentication.

Recurrent models have also been benchmarked against other architectures for combined keystroke and swipe-based mobile identification. In a dataset with 31 subjects, multimodal touch and keystroke features were fused at score level and used to train several classifiers, where RNNs achieved the best classification performance in both unimodal and multimodal configurations. Late fusion with a weighted product rule over all features yielded 95.29% accuracy and 1.78% EER, slightly outperforming a CNN baseline. A further contribution, TypeNet, used an LSTM RNN architecture trained on free-text mobile keystroke data from 26,000 participants, optimizing three different loss functions and attaining 9.2% EER with a triplet loss when balancing enrollment effort and performance [40]. Hybrid schemes that combine deep and traditional learners have also been explored: in one case, weighted outputs from an RNN and an SVM were used to authenticate ten user keystroke profiles that included motion, touchscreen, and temporal data, achieving 93.9% accuracy when motion and touchscreen modalities were fused.

CNN and transformer-style architectures represent another deep learning strand for keystroke dynamics. A transformer-based convolutional model trained on the Aalto mobile keystroke database, which consists of large-scale free-text keystroke data from approximately 260,000 subjects with five timestamp-based features per keypress, outperformed the RNN TypeNet baseline: with ten enrollment sessions, the transformer reached 3.15% EER compared to 8.00% EER for TypeNet on the same dataset [23]. Additional work has used deep neural networks constructed from stacked restricted Boltzmann machines for classification, illustrating that unsupervised pretraining can be employed in keystroke authentication pipelines [37]. Together, these studies document a clear shift from handcrafted timing features with shallow classifiers toward end-to-end or semi-end-to-end deep temporal models that better exploit large behavioral datasets. Table 3 consolidates the key behavioral biometric studies reviewed in this section, reporting modality, learning method, dataset, and performance metrics to facilitate direct comparison across approaches.

Deep learning is similarly prominent in other biometric domains relevant to continuous authentication, particularly face presentation attack detection. A systematic PRISMA-based review of AI-driven face liveness detection reports the use of convolutional neural networks, transfer learning, domain adaptation, domain generalization, generative adversarial networks, reinforcement learning, and explainable AI for detecting 2D and 3D spoofing attacks. Existing approaches rely heavily on both handcrafted features and deep feature extraction; handcrafted methods can be more robust to unknown data samples, whereas deep neural stacks offer strong performance when training and testing distributions are similar. However, current deep models generally lack generalization to invisible or synthetic attacks, such as those based on artificially generated faces, and struggle with cross-domain robustness, underscoring the need for domain adaptation and generalization strategies as future research directions. Within this face anti-spoofing context, combining multiple deep and non-deep algorithms is proposed as a way to improve resilience to unseen attacks. The review suggests that fusing handcrafted and auto-extracted deep features may yield promising performance for unknown synthetic attacks and advocates

exploring ensembles of deep models, potentially trained in a federated manner or combined through decision fusion, to increase robustness further. It also highlights a gap in “performance metrics for explanations” for explainable AI applied to biometrics, arguing that accuracy alone is insufficient when assessing deep models that must justify their decisions in security critical applications [16]. One recent line of work proposes probabilistic, kernel density estimation-based loss functions that represent each training class by a per-batch probability density and thereby replace arbitrary distance metrics with class likelihoods; this approach appears to improve discrimination of learned embeddings and has shown gains across CNN, RNN and hybrid networks in ASVspoof evaluations, with joint optimization of per-class bandwidths yielding further robustness [64]. Comparable work in fingerprint presentation attack detection has applied patch-based deep models built from Restricted Boltzmann Machines and Deep Boltzmann Machines to learn rich texture representations and achieved competitive results on LivDet benchmarks; yet these DBM/RBM pipelines appear to generalize poorly to unknown fabrication materials and can be computationally costly during training and feature extraction [65]. Such observations imply that, across different biometric modalities, measured gains in accuracy from deep feature learning may be accompanied by losses in cross-attack generalization and practical efficiency, suggesting the value of lighter models or hybrid fusion schemes for mobile deployment. Beyond keystroke and facial biometrics, deep networks appear in motion-based and multimodal behavioral authentication. A broad survey of continuous authentication and user profiling distinguishes ML and DL as core AI tools and notes that DL can leverage unsupervised learning to analyze unlabeled datasets, learning hierarchical representations from large-scale biometric data. CNNs are described as feed forward networks transforming three-dimensional input volumes into three-dimensional outputs through stacks of parameterized layers, while RNNs, including LSTMs, are highlighted for their strengths with sequential data and for supporting both supervised and unsupervised tasks by predicting future sequences from past samples. Deep neural networks with many hidden layers serve as general purpose architectures across modalities, whereas traditional algorithms like SVM and Random Forest remain common baselines for comparison [4,37]. In adjacent security domains, deep learning is intertwined with privacy-preserving and distributed training paradigms. For intrusion detection in UAV and IoT environments, stacked broad learning systems, LSTM networks, and Random Forest models have been integrated into federated and edge-based frameworks, indicating that deep models can be deployed in distributed settings where data remain local and only model parameters or gradients are shared. However, these works emphasize that combining deep methods with additional secure and privacy-preserving techniques is advisable to further enhance attack detection accuracy and to mitigate data exposure risks [17,66]. Taken together, these strands show that deep learning in continuous and behavioral biometrics spans RNN and transformer architectures for keystroke dynamics, CNN-based face anti-spoofing, and multimodal and distributed intrusion detection. Across these applications, evidence points to improved accuracy and lower EER relative to classical baselines, while also revealing open challenges in generalization to unseen attacks, explainability, and integration with privacy-preserving mechanisms in mobile and cloud-assisted environments [4,16,37,66]. Table 4 summarizes the principal multimodal fusion strategies employed in continuous authentication systems, contrasting sensor-level, feature-level, score-level, and decision-level approaches with emerging deep learning and context-aware fusion methods, along with reported performance metrics from the reviewed literature. As with Tables 1 and 2, the Accu./Compl./Priv. columns are qualitative, ordinal judgments synthesized across fusion strategies rather than scores computed on a shared benchmark; where a row cites specific quantitative results (e.g., 95.29% accuracy, 99.98% TAR), those figures come from a

single study's reported setup and should not be read as representative of the fusion-level category as a whole.

Table 4. Comparison of multimodal fusion strategies for continuous authentication.

Fusion Level	Description	Accu.	Compl.	Priv.	Example Systems/Studies	Ref.
Sensor-level Fusion	Raw signals from multiple sensors combined before feature extraction	Medium	Low	High	Accelerometer + gyroscope + magnetometer streams	[2]
Feature-level Fusion	Feature vectors from different modalities concatenated into one input	High	Medium	Medium	Keystroke + touch + motion feature concatenation	—
Score-level Fusion	Matching scores from separate classifiers combined (weighted sum, product rule)	High	Medium	Low	RNN + SVM weighted fusion; 95.29% accuracy, 1.78% EER (31 subjects)	n/a
Decision-level Fusion	Binary decisions from independent classifiers combined via voting/ AND/OR logic	Medium	Low	Low	Majority voting across keystroke, gait, context classifiers	n/a
Multimodal Fusion (LSTM)	Deep recurrent model learns joint representation from all modalities end-to-end	Very High	High	Medium	End-to-end joint representation learning across sensor streams	—
Context-aware Adaptive Fusion	System selects active modalities based on current user activity and context	High	High	Medium	GPS + app usage + keystroke; adapts strictness by location and time	[49]
Federated Fusion	Local models trained per-device; only aggregated weights shared with server	Medium–High	High	Low	eHealth IoMT federated auth; preserves device privacy	[8]

Note: EER—Equal Error Rate. Performance figures are drawn from cited studies and may vary by dataset. Rows marked “n/a” in the Ref. column (score-level RNN+SVM fusion; decision-level majority voting) correspond to configurations discussed in the prose of Section 2.6.2 whose original sources could not be verified in this draft; they are flagged explicitly rather than assigned a placeholder reference, and should be resolved by tracing each figure to its original source before submission. The Feature-level and Multimodal Fusion (LSTM) rows above describe the fusion strategy generically, without citing a specific example study, since the previously listed example figures could not be traced to a source within this review's corpus.

2.6.3. Emerging Approaches

Recent developments in continuous authentication increasingly combine advanced representation learning with privacy-preserving and distributed training paradigms. In WiFi Channel State Information-based sensing, the *IdentiFi* framework exemplifies a shift toward self-supervised and semi-supervised contrastive learning for device-free human identification. *IdentiFi* learns discriminative identity representations from unlabeled CSI time-series via customized augmentations and two contrastive modules: a temporal contrasting module that enforces cross-view prediction of future features to capture stable motion dynamics, and a contextual contrasting module that tightens intra-identity clusters while enlarging inter-identity separation in feature space. A subsequent class aware contrastive phase fine tunes the encoder with a small labeled subset and pseudo labels for remaining data, enabling high accuracy in multi-occupant environments with limited annotations and under varying wireless propagation conditions. Because the system operates purely on CSI rather than visual inputs, it provides non-intrusive, privacy-preserving identification suitable for smart homes, healthcare, and security applications in cluttered indoor environments [67]. Emerging work on behavioral biometrics also addresses adversarial robustness of motion-based continuous verification. A two-step motion biometric scheme using smartphone and smartwatch accelerometer and gyroscope measurements introduces a continuous and adaptive user verification algorithm designed for integration with existing multi-factor authentication. This approach explicitly examines susceptibility to adversarial examples generated by Zero-Order Optimization attacks and formulates a defense strategy to protect the model, motivated by evidence that adversarially crafted inputs to machine learning and deep learning classifiers can induce high misclassification rates. The same study situates motion-based authentication within a broader transition

from knowledge-based factors to behavioral biometrics and continuous authentication, and stresses the need to balance security, usability, and the temporal variability of user habits when designing adaptive verification pipelines [14]. Generative modeling emerges as both a threat and a methodological frontier. In electroencephalogram-based biometrics, Generative Adversarial Networks have been successfully used to synthesize fake EEG signals that match the statistics of genuine recordings, significantly lowering the realism barrier for attacks against EEG-based authentication. Traditional autoregressive and statistical simulation methods often show poor generalization, whereas GAN-based generators can produce high quality impostor signals that challenge the long standing assumption that EEG inherently guarantees liveness because it can only be measured from a living subject. This development exposes the absence of EEG specific liveness tests and suggests that future systems must incorporate stimulus-based or reflection-based challenge-response protocols to discriminate machine-generated from human EEG, although no concrete EEG tailored liveness scheme has yet been proposed [68]. Adversarial learning also appears in face presentation attack detection, where researchers investigate AI-based defenses against spoofing in robust face recognition. A systematic PRISMA-based review reports that convolutional neural networks, transfer learning, domain adaptation and generalization, GANs, reinforcement learning, and explainable AI all contribute to current anti-spoofing pipelines. Handcrafted texture, motion, and depth features remain more robust than pure deep features to unseen attacks, while deep neural stacks excel when train and test distributions are similar. Combining handcrafted and auto-extracted deep features is proposed as a promising direction for resisting synthetic face attacks, including those generated adversarially. Future work identified for this domain explicitly includes adversarial learning for adaptive data augmentation, meta learning for few shot adaptation to novel spoof types, and federated learning architectures such as FedPAD that allow multiple data owners to collaboratively train a global face anti-spoofing model without sharing raw data, although model-level privacy in federated settings remains unresolved [16]. Broader continuous-authentication surveys synthesize these strands into a forward-looking agenda. Proposed research directions include exploring new biometric modalities such as vein patterns and EEG signals, deploying advanced machine learning algorithms for feature extraction and classification, and integrating blockchain-based mechanisms for secure and transparent authentication. Context-aware techniques that adapt thresholds and mechanisms using factors like location and activity are seen as key to enhancing both security and usability, and multimodal fusion of physiological and behavioral traits is highlighted as a path to more robust frameworks. At the same time, open issues persist around scalability, standardized datasets and metrics, and resilience to adversarial manipulation, underscoring that future continuous authentication systems must combine sophisticated learning architectures with cryptographic protection and rigorous evaluation on realistic data [4,37].

Synthesis across ML approaches. Two conflicting narratives run through this section rather than one. Handcrafted or classical features (autoregressive EEG models, texture-based face features) are repeatedly reported as more robust to unseen or unknown attacks, while deep architectures are repeatedly reported as superior when train and test distributions match [16,68]; the literature synthesized here does not resolve which property matters more for real-world deployment, where attack distributions are by definition not known in advance. A recurring limitation across the emerging approaches (CSI-based identification, motion adversarial defenses, GAN-based EEG synthesis, face anti-spoofing) is that each is evaluated against a specific, narrow attack or condition rather than a shared adversarial benchmark, making it impossible to rank these methods against one another on a common basis. The clearest methodological trend is the movement of generative modeling from a threat (GAN-synthesized EEG and face spoofs) to a candidate defense (GAN-based adap-

tive data augmentation, federated GAN training), often within the same research area; this dual role of generative models is a distinctive feature of the recent literature that a simple modality-by-modality summary would not surface.

2.7. Security Threats and Defenses

2.7.1. Vulnerabilities of Biometric Systems

Biometric authentication mechanisms, whether physiological or behavioral, inherit a wide range of vulnerabilities that arise from both modality-specific properties and the machine learning pipelines used to implement them. Security analyses of continuous schemes stress that no modality achieves perfect accuracy, and that non-zero False Acceptance Rate and False Rejection Rate inevitably create opportunities for false matches and unauthorized access. Physiological methods often attain higher accuracy than behavioral and context-aware modes, yet they still remain susceptible to replay-style attacks such as forging fingerprints, alongside other presentation and template level threats [4,9]. Table 5 summarizes the principal attack types identified in the reviewed literature, their target modalities, severity, and corresponding defense strategies. As with Table 1, the Severity and Effectiveness columns are qualitative, ordinal judgments synthesized across the cited literature rather than scores computed by this review’s own quality-assessment rubric; where a specific attack’s impact is quantified in the reviewed studies (for example, the EER and FAR figures discussed immediately below this table), that quantitative figure should be treated as authoritative, and the table’s qualitative label read only as a relative, cross-attack comparison rather than an absolute risk measure.

Table 5. Attack types, target modalities, and defense methods in biometric authentication.

Attack Type	Target Modality	Severity	Defense Method	Effectiveness
Shoulder Surfing	Password/PIN	High	Randomized keypad, Behavioral biometrics	Medium
Smudge Attack	PIN/Pattern lock	High	Gesture-based auth, Randomized display	Partial
Phishing/Spoofing	Password, Voice, Face	Very High	Multi-factor auth, Liveness detection	Medium
Mimicry Attack	Keystroke, Gait	Medium	Multimodal fusion, Adaptive thresholds	Medium
Adversarial Attack	Touch dynamics, Keystroke	Very High	Adversarial training, Ensemble models	Medium
Replay Attack	Voice, Sensor traces	High	Challenge-response, Liveness detection	High
Model Inversion	ML models (SVM, kNN)	High	Federated learning, ZKP, Bloom filters	High
Membership Inference	Cloud-based auth models	Medium	Differential privacy, Secure aggregation	Medium
Insider Threat	Templates, Behavioral profiles	Very High	ZKP, Cancelable biometrics, Federated learning	High
ECG Template Mapping	Cardiac biometrics	Very High	Template encryption, Cancelable biometrics	Medium

Note: Severity and Effectiveness ratings are based on empirical results reported in the reviewed literature [4,9,69], synthesized qualitatively across attack types rather than computed from a single quantitative source; see the discussion following this table for the specific attack success rates underlying each rating. Red text denotes High severity, and bold red text denotes Very High severity. ZKP—Zero-Knowledge Proof.

Behavioral biometrics face particularly severe adversarial risks. Empirical evidence shows that touch dynamics cannot withstand adversarial generative attacks: when an attacker manipulates the training process, Equal Error Rate can increase from 5% to 50%. Keystroke dynamics displays even more dramatic degradation under similar generative attacks, with EER rising from 28% to 84%. Separate smartphone experiments demon-

strate that keystroke-based schemes also fail against mimicry attacks, where impostors imitate a target's typing style. Gait-based authentication is likewise vulnerable to imitation; treadmill-assisted mimicry of gait patterns can elevate False Acceptance Rate from 5.8% to 43.66%. Cardiac-based biometrics are not immune either, since attacks that map an attacker's ECG into a victim's template using a short stolen segment have achieved average success rates between 90% and 96% on datasets from 52 users [4,9]. Machine learning infrastructure itself constitutes an additional attack surface for biometric systems. Continuous authentication deployments frequently rely on classifiers such as Support Vector Machines and k Nearest Neighbors trained on physiological and behavioral data, often in cloud-assisted settings. An investigative study documents that ML approaches are vulnerable to sample inference, single and multi-sample reconstruction, and label distribution estimation attacks, while further work demonstrates that membership inference, model inversion, data deanonymization, and model extraction can all be mounted against trained models. For SVM- and kNN-based continuous schemes in the cloud, it has been shown that actual user samples are stored in authentication profiles and can be reconstructed from positive examples, enabling re identification of individuals from their biometric traces [9]. Privacy risks flow directly from these technical weaknesses. Physiological biometric templates can expose identity, health, and other biological information, as in face-based continuous authentication that collects and stores facial features which may be analyzed for emotional states. Behavioral modalities such as gait recognition, stylometry, and touch dynamics operate on daily routine data and current activities, and compromised profiles reveal behavioral patterns that cannot be revoked like passwords. Keystroke dynamics can further leak demographic attributes, including age, gender, and handedness, so successful attacks against keystroke-based systems threaten both account security and sensitive personal characteristics. Continuous monitoring and centralized storage of these biometric streams magnify the impact of any successful inference, reconstruction, or template compromise [4,9]. Context-aware and multimodal schemes introduce distinct vulnerabilities that intersect with biometric weaknesses. Systems that rely on GPS and other contextual signals may suffer from inaccurate or manipulated data, creating avenues for spoofing through falsified location or environment inputs. At the same time, capturing context such as location, browsing behavior, and application usage on a continuous basis raises data privacy concerns, since users may be unaware of the scope of contextual logging and its role in authentication decisions. Ethical analyses accordingly call for strict attention to consent, data usage, and potential bias when collecting and processing both biometric and contextual information in continuous frameworks [4,9]. Emerging machine learning paradigms introduce further systemic vulnerabilities. Federated learning in edge computing environments shares models and parameters among many clients, which opens new channels for Byzantine and poisoning attacks. In this setting, fully trusted nodes may turn rogue after passing authentication and inject malicious updates, while data and model poisoning attacks target the training process itself by sending falsified parameters or corrupt samples to influence the global model. Studies report that model poisoning can be more effective than data poisoning, especially when many clients participate in large-scale federated products, and that edge FL must account for these threats in order to maintain secure and privacy protected biometric analysis [69]. Voice- and speech-based biometric systems also face advanced adversarial scenarios. White box threat models for speaker recognition assume that attackers can access pre trained models, estimate room impulse responses, and generate adversarial audio that remains effective over the air. By simulating the acoustic channel during adversarial example generation and optimizing targeted or untargeted perturbations using gradients of the cross-entropy loss, adversaries can deceive microphone side speaker recognition systems even when signals pass through

loudspeaker and room acoustics. Such attacks demonstrate that knowledge of model internals and the physical environment suffices to construct robust adversarial audio capable of bypassing recognition, highlighting the fragility of speech-based biometrics under adaptive attack [70]. Finally, adversarial manipulation and poisoning attacks in distributed learning for UAV and IoT intrusion detection underscore broader weaknesses that are also relevant for biometric authentication. In federated and edge-based IDS for UAV networks, collaborative learning can be undermined by Byzantine nodes and poisoning strategies that degrade detection accuracy. Reviews recommend combining multiple secure and privacy-preserving approaches, rather than relying on single DNN or Random Forest models, to improve resilience. This need for ensembles and defensive training parallels challenges in biometric systems, where single classifier pipelines have repeatedly been shown vulnerable to adversarial and generative manipulation, membership inference, and model extraction [4,9,66,69].

Work in biometric presentation-attack detection further illustrates that algorithmic choices and training practices matter for operational exposure: a serial, two-stage approach built around modified MobileNetV2 backbones and aggressive data augmentation, trained on a large and diverse set of bona fide and attack images, has been reported to reach very low operational error rates (for example, a BPCER near 0.46% and EER values around 0.33%), outcomes which may reduce the frequency of fallback or re-check transactions that tend to increase the volume of biometric material routed to servers [57]. At the same time, those performance improvements appear to depend on centralized access to sizable, curated datasets and on careful weight handling, which could concentrate sensitive templates or model updates unless paired with cancelable-template schemes, cryptographic proofs, or federated training protocols that explicitly limit insider visibility.

2.7.2. Authentication Errors

Authentication decisions in biometric systems are intrinsically probabilistic, so misclassifications of both impostors and genuine users must be quantified and controlled. Continuous and behavioral schemes, in particular, rely heavily on error metrics such as False Acceptance Rate (FAR), False Rejection Rate (FRR), Equal Error Rate (EER), Half Total Error Rate (HTER), accuracy, and F1 score for performance assessment and tuning. These measures capture distinct aspects of system behavior and directly influence perceived security and usability in practical deployments [4,10,11]. Formally, FAR denotes the proportion of illegitimate attempts that are wrongly accepted. It can be expressed as the percentage of falsely accepted illegitimate users relative to the total number of impostors accessing the system, as in

$$\text{FAR} = \frac{\text{Number of falsely accepted illegitimate users}}{\text{Total number of imposters}} \times 100\% \quad (1)$$

and more generally as the ratio of false positives to the sum of true negatives and false positives [4,11]. FRR, in contrast, captures the proportion of legitimate attempts that are incorrectly rejected, defined either as the percentage of falsely denied legitimate users over all genuine users or as the ratio of false negatives to the sum of true positives and false negatives [4,11]. HTER aggregates these two error components as the simple average

$$\text{HTER} = \frac{\text{FAR} + \text{FRR}}{2} \% \quad (2)$$

which is also referred to as balanced accuracy in some keystroke dynamics studies. At the operating point where FAR and FRR are numerically equal, the system attains the EER, a widely used scalar summary for biometric authentication performance. Adjusting the

decision threshold directly shapes this error trade-off. Increasing the threshold reduces FRR while increasing FAR, whereas decreasing the threshold lowers FAR (and thus nominally strengthens security) at the cost of higher FRR, so threshold choice must be carefully optimized. Industrial practice introduces additional constraints; for example, European access control standards specify that FAR should be below 1% and FRR no more than 0.001%, which are stringent targets for continuous, sensor-based schemes. In keystroke dynamics research, EER is the most common primary metric, with several recent works using EER alone, and others combining EER with accuracy, or explicitly reporting FAR and FRR alongside F1 score and HTER to characterize classifier behavior. This diversity of metrics has been adopted broadly across user authentication models and is accepted in both academic and industrial evaluations [11].

Continuous-authentication surveys reiterate that these error measures are central not only for benchmarking algorithms but also for understanding practical viability. Accuracy is defined as the percentage of correctly matched authorized users among all users, while FAR and FRR quantify, respectively, the likelihood of mistakenly admitting unauthorized users and mistakenly rejecting authorized users. In this framing, FAR and FRR are treated as core design levers: high FAR is associated with increased security risk, whereas elevated FRR undermines usability and user acceptance. Continuous User Biometric Authentication Systems therefore face an inherent trade-off between these error types, and usability discussions explicitly stress that system design must prioritize acceptable FRR levels without allowing FAR to rise to dangerous values [4,10].

Error reporting in broader behavioral biometric research reflects similar concerns. A systematic review of behavioral biometric authentication notes that EER, accuracy, FAR, and FRR are the most popular model evaluation metrics, yet also finds inconsistent terminology, such as “Correct Recognition Rate” in place of accuracy and “False Positive Rate” for FAR. This lack of standardization hinders comparison across studies and complicates meta-analytic assessment of authentication error behavior. In addition, study quality assessments reveal that many works do not adequately describe experimental design, sampling, or evaluation procedures, further limiting the reliability and reproducibility of reported FAR, FRR, and EER values [10].

In continuous authentication, security analyses connect these error metrics directly to vulnerability. One line of work characterizes security by performance measures such as accuracy, FAR, and FRR, together with the ease of forging a given modality and the range of attacks, including mimicry and template leakage. Since neither physiological nor behavioral mechanisms achieve perfect accuracy, non-zero FAR and FRR translate into concrete probabilities of false matches and missed detections. Evidence indicates that physiological methods typically achieve better accuracy than behavioral and context-aware modes, implying comparatively lower EER and FAR, although they still remain subject to replay-style attacks such as fingerprint forging. Behavioral approaches, especially those based on touch dynamics and keystroke dynamics, can exhibit relatively high FAR and FRR, particularly when evaluated on small datasets, reinforcing concerns that error rates may remain too large for high-assurance deployments without additional safeguards [9]. Several experimental RNN-based studies report a wide spread in error figures that appears closely linked to modality and dataset choice: inertial gait models achieved EERs near 11.48% (training) and 7.55% (testing) on OUIDB, handwriting and signature methods report EERs around 3.8–6.44%, and mouse- or keyboard-based approaches sometimes show single-digit EER or very high accuracy on small, controlled collections [54]. Such variance may suggest that low scalar metrics can be dataset-specific and should be interpreted alongside experimental details like participant count, session variability, and attack simulations rather than taken as definitive evidence of readiness for deployment.

Beyond scalar metrics, continuous-authentication surveys also call for expanded evaluation frameworks that relate error behavior to usability, scalability, and robustness. Existing research predominantly focuses on supervised learning and on FAR and FRR, while identifying research gaps in comparative analysis across biometric categories and in using richer metrics for practical success factors. Future directions proposed for continuous authentication include developing standardized benchmark datasets and uniform evaluation protocols, exploring additional error related measures, and integrating context-aware mechanisms that adapt thresholds dynamically based on environmental factors and user activity, all with the goal of managing FAR and FRR in a way that better fits real-world operation [4].

Synthesis across error-reporting practice. This section surfaces a conflict that individual study summaries obscure: physiological modalities are reported as achieving lower EER/FAR than behavioral modalities [9], yet the RNN-based comparison above shows behavioral EERs spanning roughly 3.8% to over 11% depending on modality and dataset [54], a range wide enough that some behavioral configurations outperform typical physiological baselines while others do not; performance ranking by modality category is therefore not stable across the corpus and depends heavily on which specific dataset and protocol is used. A recurring limitation, flagged independently in [10,11], is inconsistent metric terminology (e.g., “Correct Recognition Rate” for accuracy, “False Positive Rate” for FAR) and incomplete reporting of sampling and evaluation procedure, which together mean that a headline EER figure alone cannot be safely compared across the studies reviewed without checking the underlying protocol. The common methodological trend is a field-wide reliance on scalar operating-point metrics (FAR, FRR, EER) computed on small, single-dataset evaluations, with calls for standardized benchmarks appearing across nearly every source cited in this section but no evidence, within the corpus reviewed here, that such a shared benchmark has yet been adopted in practice.

2.8. Future Directions

2.8.1. Integration with Wearables and IoT

Continuous authentication in pervasive environments increasingly depends on distributed infrastructures where sensing, learning, and access control span constrained devices, edge nodes, and cloud services. Fog and IoT cloud architectures decentralize data processing, access control, application security, and network analytics across edge networks to guarantee data integrity and security, while intelligent gateways handle privacy sensitive information generated by sustainable architectures and transmit it toward appropriate edge servers. In such settings, end user authentication is explicitly identified as a core capability of fog-based services used to secure communication over adversarial networking devices, with user modeling trained on real time data streams from IoT devices. To satisfy delay and mobility requirements, network infrastructures prefer fog computing, yet privacy sensitivity of statistical databases prevents central storage of raw user data for predictive modeling, which motivates federated learning as a distributed machine learning paradigm for training authentication and analytic models on sensitive IoT data without centralized aggregation [8]. Within this edge-centric perspective, surveys on federated learning in edge computing classify a broad range of prior work that addresses edge architectures, offloading, and management of wireless and mobile computing resources, and conclude that dedicated systematic reviews on federated learning for edge computing remain scarce, particularly with respect to hardware requirement challenges [69].

Wearable devices and mobile sensors act as primary sources of behavioral and contextual data for continuous authentication in these IoT cloud environments. Smartphone-based Continuous User Biometric Authentication Systems already encounter notable energy con-

straints, since sensors such as light, proximity, gyroscope, and digital compass consume substantial battery power and sensory data collection itself increases energy usage, while CUBA requires sensors and processing pipelines to remain powered throughout full active sessions [4]. Extending continuous authentication to wearables amplifies these constraints due to even more limited battery capacity and computation, which strengthens the case for offloading heavy learning tasks to fog or cloud infrastructures while keeping raw biometric data local through federated training. Practical implementations indicate that an end-to-end pipeline can stream touch and motion traces from a mobile client to server-side models for real-time classification; one such mobile banking deployment reported session-level true positive rates approaching 99% after periodic model updates, while observing little additional power or memory overhead from the authentication layer [53]. Their findings also suggest that aggregating decisions at the session level and updating models with new data are pragmatic ways to counter drift in behavioral signals without moving costly computation onto constrained devices. In healthcare-oriented wearable deployments, gait-based identification systems designed specifically for elderly users demonstrate that inertial sensor signals can support reliable continuous identity recognition with low computational overhead on wearable devices, offering a practical authentication layer for remote patient monitoring environments [71]. In smart healthcare and eHealth scenarios, data cloud-centric architectures explicitly employ federated learning to train predictive models over sensitive sensor readings transmitted from mobile sinks and edge IoT devices to fog and cloud centers, thereby aiming to preserve privacy while supporting scalable analytics and authentication related services [8]. These designs illustrate how integration with IoT requires joint consideration of network latency, mobility, storage limits, and privacy, rather than treating continuous authentication as a purely device local problem.

Trust management and message authentication in IoT networks further shape how continuous authentication can be embedded into wearable and sensor-rich ecosystems. Trust-based monitoring schemes with intelligent middleware agents use modified learning processes in which hidden layer outputs of a machine learning model are verified against knowledge bases and agent lists to identify imitative device behavior, decreasing misdetection risk by distributing verification across middleware components rather than concentrating checks in a single layer [7]. Only devices that qualify under this learning process are permitted to employ the message authentication scheme orchestrated by the middleware manager, and agents continuously update authentication lists based on new and existing device communications. Such architectures align naturally with continuous authentication based on behavioral biometrics emitted by wearables and smartphones, since the same distributed agents can evaluate both trust in device behavior and user identity signals, while also balancing communication cost and energy consumption across the network [7,8].

Blockchain-based smart home frameworks provide another avenue for integrating continuous authentication into IoT environments. Attribute-based access control mechanisms implemented via Ethereum smart contracts manage policies and attributes for subjects and objects and perform access control for IoT systems, offering an unforgeable and auditable foundation for preventing unauthorized data access and protecting information security. Additional blockchain proposals address data privacy by exploiting public key infrastructures, differential privacy techniques, and decentralized consensus mechanisms such as Proof of Authority to manage home appliances in a distributed manner, although challenges remain regarding deanonymization attacks, central administration points, and transaction latency [60]. These works suggest that wearable-derived behavioral biometrics could be fused with attribute-based and token-based access control in smart homes, where edge gateways and blockchain smart contracts jointly authenticate users

and devices while maintaining tamper resistant logs of access decisions. Some authors argue that traditional brokered models fail to give users fine-grained control over personal data and that decentralized ledger technologies, combined with smart contracts and distributed storage, may enable more user-centric, auditable transactions in Ambient Assisted Living and smart home settings [72]. That said, practical constraints such as throughput limits and governance of administrative nodes appear to remain unresolved, so DLT-based schemes seem to trade certain centralization risks for gains in transparency and user agency. Continuous-authentication surveys identify integration of context-aware techniques, multimodal biometrics, and advanced machine learning as explicit future directions, including exploration of novel modalities such as motion-based authentication and the fusion of physiological and behavioral traits to adapt authentication to diverse user scenarios [4]. Within IoT and wearable rich settings, this implies combining motion and activity signatures captured by accelerometers and gyroscopes with contextual attributes such as location, time, and application usage that are already monitored by connected devices [4,8]. Privacy-preserving mechanisms are simultaneously highlighted as essential for such pervasive deployments: zero-knowledge-proof-based protocols can prevent servers from learning cryptographic keys, metadata, or explicit authentication information, Bloom filter-based constructions can store irreversible and unlinkable biometric templates, and cancelable biometrics can enable revocation and re issuance of biometric identifiers if compromise occurs, including compromise by curious or malicious insiders in cloud-based continuous authentication backends [4,9].

Ethical and usability considerations remain integral in extending continuous authentication across wearables and IoT. Research on behavioral biometric continuous authentication adoption shows that privacy concerns and perceived response costs, such as fears of biometric misuse, influence intention to use BBCA technologies, while self-efficacy, perceived response efficacy, and innovativeness act as facilitators of adoption [1]. These findings suggest that large-scale IoT deployments involving continuous monitoring via wearables and ambient sensors must explicitly address user perceptions of privacy risk, trustworthiness of authentication agents, and clarity of consent, especially when behavioral data are processed in distributed fog and federated learning architectures rather than solely on device [1,4,8].

2.8.2. Blockchain for Identity Management

Identity management in highly connected environments increasingly leverages blockchain as a foundational security primitive for authentication, integrity, and privacy preservation. In V2X-focused 6G scenarios, blockchain technology is proposed to provide a tamper resistant, distributed ledger that secures sensitive data and prevents unauthorized modification, while supporting decentralized trust among vehicles, infrastructure, and users [18]. Each vehicle or infrastructure node can be assigned a unique digital identity recorded on-chain, enabling authenticated interactions without relying on a single centralized authority. This use of immutable records is framed as particularly suitable for safeguarding data exchanged in latency sensitive V2X communication, where traditional centralized identity services may struggle to meet both availability and security requirements [6,18].

Several works extend this identity-centric view of blockchain into broader cyber-physical and virtual environments. In smart home ecosystems, attribute-based access control and privacy preservation are implemented via Ethereum smart contracts, with each participant holding a distinct public-private key pair and blockchain address. Smart contracts manage registration of users and devices, maintain access policies, and enforce authentication decisions in a decentralized manner. To protect user privacy during distributed machine learning over smart home data, differential privacy-based perturbation of

local gradients is combined with blockchain logging, which both hardens records against tampering and aims to preserve confidentiality of model updates [60]. Parallel work on Metaverse security explicitly argues that blockchain-based identity verification can underpin a zero-trust architecture, providing decentralized identity management and data integrity guarantees for virtual environments where assets and interactions are highly dynamic [73].

Mobile contexts motivate dedicated blockchain-based identity schemes. A mobile chain framework proposes blockchain backed authentication for mobile environments, implemented on Ethereum and designed to offer anonymity, untraceability, confidentiality, integrity, and decentralization simultaneously. Formal verification using AVISPA is reported, and the authors claim that mobile chain outperforms existing mobility networks in terms of exposure to security threats, while remaining computationally efficient enough for resource-constrained wireless and mobile scenarios. In the same survey, anonymous authentication schemes based on zero-knowledge proofs are discussed as privacy-preserving cryptographic tools that allow users to authenticate repeatedly without revealing their identities, highlighting that blockchain-integrated identity systems can embed zk proof mechanisms such as zkSaaS and zkLogin to outsource proof generation and bind on-chain identities to off-chain OpenID credentials without exposing the linkage [2].

Identity and access control in V2X and 6G environments are another active focus area. Analyses of 6G-enabled V2X architectures describe blockchain as an enabler for secure and efficient management of massive data generated by heterogeneous nodes, emphasizing its role in providing authentication and integrity through immutability and distributed storage [6,18]. One proposal defines a blockchain-based V2X solution in which each vehicle and infrastructure node possesses a unique digital identity recorded on-chain, and blockchain-mediated communication ensures that data are exchanged only among authorized entities, thereby improving overall trust in the system. A complementary line of work explores joint use of blockchain and federated learning in V2X 6G networks, with blockchain responsible for guaranteeing data integrity, authentication, and tamper resistance, while federated learning adjusts blockchain parameters based on features such as number of vehicles, threat levels, and available resources. This combination is argued to facilitate interoperability between different blockchain networks and to support identity management across dynamic, heterogeneous 6G environments [18].

Privacy-preserving mechanisms integrated with blockchain-based identity are receiving particular attention. Differential privacy has been combined with stochastic gradient descent in decentralized learning frameworks like LearningChain, where perturbed gradient updates are committed to a public blockchain, claiming both reduced error rates and resistance to byzantine attacks via l nearest aggregation. Additional studies incorporate repeated additive differential privacy noise into blockchain transactions to strengthen user trust in distributed machine learning and to defend against linkage and double-spending-style attacks, suggesting that such perturbation can make individual records indistinguishable from their neighbors while maintaining acceptable utility. In smart grid scenarios, blockchain-based token banks and de-regulated micro-grid architectures employ Laplace noise and enhanced proof of authority consensus to hide consumption patterns and enhance user privacy, while still supporting real time trading and energy management; identity in these settings is thus mediated through blockchain addresses whose linkage to real-world entities is shielded by differential privacy [60].

Emerging virtual spaces such as the Metaverse further expand the envisioned role of blockchain in identity management. A zero-trust architecture tailored for virtual environments integrates blockchain-based decentralized identity verification with dynamic trust assessment, aiming to secure eHealth and asset management applications where user identities and cre-

dentials must be protected across multiple domains and service providers. Surveyed work on Metaverse as a Service models stresses that effective identity and access management requires strong privacy policies and scalable blockchain infrastructures, while also acknowledging that empirical validation and real-world deployment studies are still limited [73,74]. These findings, taken together with identity-centric blockchain proposals in smart homes and V2X, indicate a trajectory in which blockchain acts as a unifying substrate for identity, authentication, and auditability across mobile, IoT, and virtual ecosystems [2,18,60,73]. Moreover, integrating open-source cyber intelligence, including signals drawn from social media and web-based feeds, may help on-chain identity systems detect anomalous enrollment or credential-use patterns by providing contextual threat indicators that complement cryptographic checks. Multi-agent frameworks that aggregate translation, sentiment, topic and anomaly detection from heterogeneous text streams appear to offer a pragmatic path for this kind of external risk assessment, though coupling such pipelines to ledger records requires careful tuning to avoid privacy leaks and false alarms [75].

3. Methodology

This systematic review was not prospectively registered in a public registry such as PROSPERO; no formal protocol was published in advance of conducting the review.

3.1. Publication Bias and Duplicate Publication Handling

Two distinct risks are addressed here separately, since they are often conflated in systematic reviews of this kind. First, *publication bias* refers to the systematic underrepresentation of null or negative findings in the retrievable literature: studies reporting weak discrimination between legitimate users and impostors, or failed deployments of a proposed authentication mechanism, are less likely to be submitted or accepted for publication than studies reporting strong EER or accuracy figures. This review does not attempt a quantitative publication-bias correction (e.g., a funnel plot or Egger's test), because such tests presuppose a single pooled effect size across comparable study designs, which does not hold across the heterogeneous corpus of system papers, ML benchmarks, and user studies synthesized here. Instead, three qualitative mitigations were applied: (i) six databases spanning both computer-science-indexed (IEEE Xplore, ACM Digital Library) and multidisciplinary (Scopus, Web of Science, ScienceDirect, SpringerLink) sources were searched jointly, reducing venue-specific selection effects; (ii) backward snowballing and forward citation tracking (Section 3.5) were used to surface lower-visibility studies not well indexed by keyword search alone; and (iii) the quality-assessment rubric (Study Quality and Risk-of-Bias Assessment, below) explicitly scores each study independently of the direction or strength of its reported result, so that studies with modest performance are not systematically down-weighted relative to studies reporting strong metrics. We nonetheless note this as a limitation in Section 5.2: reported EER and accuracy figures synthesized in Section 4 should be read as drawn from a published-literature distribution that may skew optimistic relative to the true population of attempted behavioral authentication systems, including unpublished negative results from industry deployments.

Second, *duplicate publication* refers to the same underlying study or dataset appearing across multiple papers (e.g., a conference version later extended into a journal article, or the same dataset re-analyzed across two closely related publications by overlapping author teams), which is a distinct risk from simple duplicate database records of the same reference. Record-level duplicates (the same reference indexed by more than one database) were removed at the deduplication stage using Zotero with manual verification, as reported in the PRISMA flow (Section 3.5). Study-level duplicate publication was checked separately during full-text screening: for each candidate study, reviewers compared reported sample

sizes, dataset names, and author overlap against already-included studies; 100 records were excluded at the full-text stage specifically as duplicate or overlapping reports of an already-included study (Section 3.5), distinct from the 100 records excluded for insufficient methodological rigor and the 30 excluded as non-peer-reviewed. Where a conference and an extended journal version of the same system were both retrieved, only the more complete (typically the journal) version was retained in the 80-study corpus, to avoid double-counting the same empirical evidence in the thematic and metric synthesis in Section 4.

3.2. Keyword Selection and Search Criteria

For replicability, the five Boolean search strings actually used in this review, their exact execution dates (September 2025 through January 2026), and per-database record counts are given in full later in this section and in the Supplementary Materials (Search_Strategy.txt); readers seeking only to replicate the search can proceed directly to those strings. The remainder of this section situates these choices relative to comparable keyword strategies in prior systematic reviews.

Keyword formulation in this review follows established systematic review practices that start from clearly articulated research questions and then derive controlled vocabularies to ensure comprehensive yet focused retrieval. In face liveness detection, for instance, research questions are explicitly mapped to themes such as attack types, datasets, methods, challenges, and future directions, and this mapping then guides the construction of the search string. A similar logic is adopted here: primary concepts cover continuous authentication, behavioral biometrics, intelligent personal assistants, and mobile sensing, while secondary concepts capture machine learning techniques, evaluation metrics, and threat models relevant to mobile and IoT ecosystems. Search term construction benefits from prior work in adjacent domains. For AI-based face anti-spoofing, a two-phase procedure is used: first, fundamental keywords such as “Face Anti Spoofing”, “Face Liveness Detection”, “Face Presentation Attacks”, “Artificial Intelligence”, and “Domain Adaptation” are defined; second, these are combined with secondary terms including “Machine Learning”, “Deep Learning”, “Domain Generalization”, “Reinforcement Learning”, and “Face Biometric Spoofing” using Boolean operators AND/OR to generate database specific queries [16]. Intelligent personal assistant surveys similarly derive search strings from PICOC-based research questions and then constrain scope using phrases such as “intelligent personal assistant”, “intelligent personal attendant”, “automated personal assistant”, and “automated personal attendant” in the abstract field [5]. In continuous authentication profiling, keywords like “continuous user authentication”, “biometric authentication”, “behavioral authentication”, and “user identity verification” are combined with Boolean operators and filters to capture recent work [4]. These precedents inform the present review’s strategy of pairing modality and device terms (for example, keystroke, touch, motion, accelerometer, gyroscope, smartphone) with authentication and intelligent agent terms under a unified Boolean framework. Following this two-phase logic, the present review constructed five database-specific search strings combining primary and secondary terms via Boolean operators. The strings were applied to the title, abstract, and keyword fields across all queried databases:

- **String 1 (core modalities and continuous authentication):**
 (“continuous authentication” OR “continuous user authentication”)
 AND (“behavioral biometrics” OR “behavioral biometrics”) AND (“mobile device” OR “smartphone”)
- **String 2 (intelligent personal agents):**

(“intelligent personal assistant” OR “personal agent” OR “IPA”) AND
 (“authentication” OR “user verification”) AND (“mobile” OR “IoT”)

- **String 3 (biometric modalities and machine learning):**

(“keystroke dynamics” OR “gait recognition” OR “touch dynamics” OR
 “swipe gesture”) AND (“machine learning” OR “deep learning”) AND
 (“authentication”)

- **String 4 (privacy-preserving mechanisms):**

(“federated learning” OR “cancelable biometrics” OR “zero-knowledge
 proof”) AND (“biometric authentication” OR “continuous authentication”)

- **String 5 (IoT and wearable contexts):**

(“Internet of Things” OR “wearable” OR “edge computing”) AND (“behavioral biometrics” OR “continuous authentication”)

The final database search (database queries) was conducted between September 2025 and January 2026; subsequent screening, full-text eligibility assessment, and backward and forward snowballing extended the overall project timeline through April 2026. Full search strings for each individual database, including any platform-specific syntax adjustments, are provided in the Supplementary Materials (Search_Strategy.txt). Each string was executed sequentially across all six databases on a fixed date (String 1: September 2025; String 2: October 2025; String 3: November 2025; String 4: December 2025; String 5: January 2026), together yielding approximately 1744 records before deduplication. The database search strings themselves were executed sequentially between September 2025 and January 2026 as detailed above; the subsequent screening, full-text eligibility assessment, and backward/forward snowballing (Section 3.5) extended the overall project timeline through April 2026, which is the date referenced elsewhere in this manuscript as the search “conclusion” date for the purpose of indexing-lag discussion (Section 4.2).

The five strings were deliberately scoped to a constrained set of core terms and their British/American spelling variants (e.g., “behavioural”/“behavioral”) rather than an exhaustive synonym list, to preserve precision given the breadth of the five modality and infrastructure themes covered; broader synonyms (e.g., “implicit authentication”, “active authentication”, “passive authentication”, “continuous verification”, “voice assistant”) were considered but not included, and their omission may have excluded a small number of relevant records indexed only under those terms. To gauge the practical impact of this scoping decision, a supplementary, non-systematic check was performed: each of the five omitted synonym terms was run individually against Scopus and IEEE Xplore (the two highest-yield databases in the primary search), restricted to the same 2017–2025 window. This check returned 34 additional candidate records not present in the original 1744-record pool, of which 6 passed title/abstract screening and 2 passed full-text eligibility; both were studies on passive authentication using gait and touch fusion already conceptually represented in the included corpus via other terminology (e.g., “context-aware” and “transparent authentication”). These two records were not retrospectively added to the primary corpus, since doing so after quality assessment and thematic coding had already been completed would have required re-running the full screening pipeline; instead, they are noted here as evidence that the keyword scoping in Section 3.2 excluded a small, largely redundant set of records rather than a substantively different body of evidence. A second, distinct limitation concerns search scope rather than vocabulary: across all six databases, search strings were applied to title, abstract, and keyword fields only, following the field-restriction convention used in the prior systematic reviews this protocol was modeled on [4,10,16]. This restriction excludes studies in which terms central to this review, such as “intelligent personal agent”

or “behavioral biometric continuous authentication”, appear primarily in the full text rather than in the title, abstract, or keywords, and we did not perform a supplementary full-text search to recover such cases. Both limitations are stated here explicitly as constraints on recall rather than addressed retroactively, since correcting them would require re-executing the search with a revised protocol rather than reinterpreting the existing results.

Database selection is aligned with prior systematic studies on behavioral biometrics and IPAs. One scoping review targeting mobile sensor-based behavioral authentication draws from Web of Science Core Collection, Inspec in Engineering Village, Applied Science & Technology Source, IEEE Xplore, motivated by their relevance and size, and documents a combined initial yield of 22,537 studies before de-duplication [10]. For continuous-authentication surveys, IEEE Xplore, ACM Digital Library, ScienceDirect, and other major academic databases are queried using keyword sets centered on continuous and biometric authentication, with peer-reviewed status and English language as basic filters [4]. In IPA research, databases such as SpringerLink and Scopus are explicitly mentioned as primary sources to maximize the probability of capturing highly relevant work, and more than 3472 articles over a six-year period are screened before 58 primary studies are retained [5]. Guided by these examples, the current review targets multidisciplinary bibliographic platforms that jointly cover computing, signal processing, security, and human–computer interaction, and applies consistent keyword logic across them.

Search scope restriction and refinement rely on explicit temporal, linguistic, and topical constraints. In AI-driven face liveness detection, only original research articles between 2011 and 2022 that contain the fundamental or secondary keywords in the title, abstract, or full text and that address face presentation attack detection are included, while non-English, duplicate, non-full text, and thematically irrelevant papers are excluded [16]. IPA mapping limits inclusion to studies published from 2012 onward, written in English or Brazilian Portuguese, that address IPA or closely related terms in the title or abstract, and uses an intervention-based filter on “personal assistant” and “personal attendant” to remove conversational agents outside the defined scope [5]. Continuous-authentication surveys adopt comparable inclusion criteria requiring direct relevance to continuous user authentication, peer review, methodological rigor, and full-text availability, while excluding non-relevant, low-quality, and duplicate items [4]. These patterns motivate the present review’s focus on peer-reviewed, full-text articles within a recent time window, explicitly tied to continuous mobile authentication, behavioral biometrics, and intelligent personal agents.

Operationalization of these criteria follows PRISMA-style multi-stage screening. In face anti-spoofing, a three-step protocol is implemented: formulation of research questions, database searching with Boolean queries, and application of inclusion and exclusion criteria, all documented via a PRISMA flow diagram that traces identification, screening, eligibility, and inclusion. An initial 283 Scopus and 188 Web of Science records (471 total) are reduced by removing 187 duplicates, abstract-based screening, and full-text quality assessment, ultimately yielding 95 articles for detailed analysis [16]. The IPA review likewise details staged impurity removal, title and abstract screening, automatic duplicate elimination based on titles and authors, and final full-text filtering, emphasizing that many conversational agent papers must be excluded even when keyword matches occur [5]. A recent continuous-authentication survey describes a comparable procedure: keyword driven retrieval, predefined inclusion and exclusion rules, independent screening by multiple reviewers, resolution of disagreements via discussion or third reviewer, and structured data extraction from retained studies [4]. In mobile behavioral biometrics, reviewers additionally perform inter-rater reliability checks during title and abstract screening, achieving above 80% and in one case 99.9% agreement before independently processing remaining records [10]. The present work adopts these methodological safeguards by combining

staged PRISMA-based screening, duplicate removal, independent reviewer assessment, and standardized extraction forms.

Finally, search criteria are complemented by supplementary strategies aimed at completeness. IPA studies explicitly extend database queries with manual searches of reference lists from primary studies and other reviews, and include books and technical reports where appropriate, while still subjecting these to the same inclusion and quality assessment rules [5]. Continuous-authentication surveys likewise report checking reference lists and related work sections of included papers to identify additional relevant publications not captured by initial keyword queries [4]. In the current review, analogous backward snowballing on key continuous mobile authentication, behavioral biometric, and IPA articles is used to minimize the risk of missing influential contributions that may use slightly different terminologies than those specified in the primary keyword set [4,5,10,16]. For example, a recent MotionSense study reports identification accuracy above 98.4% using a hybrid CNN–RNN architecture on accelerometer, gyroscope and magnetometer data, which may suggest strong potential for motion-based modalities when paired with modern deep models [52]. Yet the same work also cautions that performance depends on the range of activities represented and does not explicitly handle abrupt changes in user behavior, a limitation we take into account when defining our temporal and topical filters.

3.3. Inclusion and Exclusion Criteria

This review applied the following explicit inclusion and exclusion criteria during title/abstract and full-text screening.

Inclusion criteria:

- Published in English, as a peer-reviewed journal article, conference paper, or review article.
- Published between 2017 and 2025.
- Indexed in at least one of the six queried databases (Scopus, Web of Science, IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink), or identified through backward or forward snowballing from an included study.
- Topically addresses continuous authentication, behavioral biometrics, or intelligent personal agents in the context of mobile devices, wearables, or IoT.
- Reports an identifiable authentication-related research objective and at least some empirical or system-level grounding, corresponding to a non-zero score on gatekeeping criteria C1 and C2 of the quality assessment rubric (Section 3.5).

Exclusion criteria:

- Not published in English.
- Not peer-reviewed (e.g., preprints, non-archival technical reports, or document types outside journal articles, conference papers, and review articles).
- Duplicate records, including near-duplicate conference and journal versions of the same study.
- No identifiable authentication-related objective or no empirical/system-level grounding (a score of 0 on gatekeeping criteria C1 or C2).
- Topically unrelated to continuous authentication, behavioral biometrics, or intelligent personal agents in a mobile, wearable, or IoT context (e.g., agent-based systems addressing unrelated application domains).

The remainder of this section situates these criteria relative to comparable choices made in prior systematic reviews on continuous authentication, behavioral biometrics, and intelligent personal assistants.

Selection of primary studies followed a structured, multi-stage process that combined keyword-based database searches with explicit relevance, quality, and language constraints. Continuous-authentication surveys on user profiling first limited inclusion to

peer-reviewed, full-text articles in English that directly address continuous user authentication, then excluded work lacking sufficient methodological rigor, non-peer-reviewed publications, and duplicates. Two independent reviewers screened titles and abstracts using a standardized checklist, resolved disagreements through discussion or third reviewer input, and then conducted a full-text assessment focused on methodology, findings, and relevance to survey objectives [4]. Comparable screening pipelines for behavioral biometric scoping reviews required that studies use built-in smartphone or tablet sensors, such as accelerometers, gyroscopes, magnetometers, and touch interfaces, explicitly for verification, detection, and authentication of device users. Some empirical studies paired built-in sensors with bespoke input designs to increase the distinctiveness of behavioral traces; for example, one keystroke dynamics study assigned each participant a per-user randomized keypad generated by a Mersenne twister, combined touch and motion features with feature selection, and compared multiple classifiers, reporting improved user classification and best results with a Manhattan distance classifier [76]. Such methodological choices may make a study easier to appraise for methodological completeness and thus more likely to meet inclusion thresholds. After initial duplicate removal across multiple databases, titles and abstracts of more than 14,000 records were screened, followed by full-text review of over 200 articles against predefined criteria; only those fully meeting the eligibility rules were extracted for detailed analysis [10].

Domain specific systematic reviews provide additional precedents for defining inclusion and exclusion rules. AI-driven face liveness detection studies restricted the corpus to original research articles published between 2011 and 2022 that contained face anti-spoofing-related keywords in the title, abstract, or full text and that proposed solutions for face presentation attack detection. Exclusions removed non-English papers, duplicates, items without full text, and all research not substantively related to face liveness or presentation attacks. A three step screening protocol combined abstract-based filtering, full-text relevance checks, and quality analysis, where articles lacking clear findings, empirical support, well presented goals and outcomes, or adequate referencing were eliminated before final inclusion [16]. Another systematic review on continuous authentication adopted similar constraints, retaining only research that answered at least one predefined research question, used biometric-based continuous authentication, and provided methodological detail sufficient for critical appraisal, while omitting low-quality, tangential, or incomplete reports. In the intelligent personal assistant domain, mapping studies applied population and intervention-based exclusion rules that removed any article not addressing IPA or related acronyms and any work that did not contain “personal assistant” or “personal attendant” (or plurals) within the relevant fields. Additional stages discarded impurities such as conference names mistakenly captured by keyword queries, filtered records by title and abstract to ensure an IPA centered focus, automatically eliminated duplicates across databases using title and author matches, and finally excluded full texts whose main subject fell outside the IPA scope. Eligible studies were limited to those treating IPA in general, including software or architectures from student and professional contexts, and had to be written in English or Brazilian Portuguese; technical reports could be included if they satisfied these criteria. A subsequent quality assessment evaluated purpose, contextualization, literature review, related work, and methodology before final acceptance [5]. PRISMA-oriented systematic reviews of face anti-spoofing and continuous authentication operationalize these rules through detailed flow diagrams that document identification, screening, eligibility, and inclusion. For example, one face liveness review retrieved 471 records from Scopus and Web of Science, removed 187 duplicates on the basis of DOI and title, then applied title- and abstract-based inclusion and exclusion criteria to 284 documents. Ninety eight records failed the inclusion criteria at this stage, leaving

186 studies for full-text eligibility checks; the final corpus comprised 95 articles that satisfied all inclusion and quality thresholds. Authors explicitly linked this staged reduction to predefined research questions and keyword strategies, emphasizing that only those works addressing face presentation attack detection with adequate empirical grounding were retained [16]. A continuous-authentication survey reports an analogous three step PRISMA-based protocol, starting from research question formulation, progressing through database querying with Boolean combinations of continuous authentication related keywords, and applying criteria centered on topical relevance, peer review, language, and methodological completeness until a stable set of high-quality studies remained [4,16].

Methodological descriptions in continuous user authentication reviews further specify how inclusion decisions were documented and controlled. After initial keyword-based retrieval from IEEE Xplore, ACM Digital Library, ScienceDirect, and related databases, articles were admitted only if they addressed continuous user authentication, provided the full text in English, and met peer review standards. Exclusion rules removed studies outside the continuous authentication domain, those with insufficient methodological rigor, and duplicates. Data extraction was then limited to this vetted subset, and inter-rater reliability checks ensured consistency of screening and coding; discrepancies between reviewers were resolved through consensus discussions, with the process periodically audited for quality assurance. Ethical safeguards, including plagiarism prevention, proper citation, and adherence to guidelines for handling extracted data, completed the criteria framework, with all data used solely for the purposes of the survey [4].

Scoping work on behavioral biometrics for mobile screen time provides quantitative evidence of how strict inclusion and exclusion criteria shape the final dataset. From an initial pool of 22,537 records aggregated across multiple databases, automated and manual duplicate removal reduced the set to 14,179 for title and abstract screening. Only 207 articles progressed to full-text retrieval, and 122 ultimately satisfied inclusion rules that required focus on built-in smartphone or tablet sensors for verification, detection, or authentication of the device user. Studies that did not employ relevant sensors, fell outside smartphone or tablet contexts, or failed to implement user verification or authentication tasks were excluded at the full-text stage [10]. Research shows that this combination of topic specific constraints, PRISMA-based staging, language and peer review filters, and explicit quality assessment criteria constitutes a robust template for defining inclusion and exclusion rules in systematic reviews of continuous authentication, behavioral biometrics, and intelligent personal assistants [4,5,10,16]. Before describing quality scoring, it is necessary to clarify what counts as a “primary study” in this corpus, since the 80 included works are not uniformly original empirical studies. Of the 80, 53 report a deployed or simulated authentication mechanism, a behavioral or physiological biometric pipeline, or an evaluated security/privacy mechanism, and are treated as primary empirical sources in the strict sense. The remaining 27 are themselves systematic reviews, narrative surveys, or scoping reviews of continuous authentication, behavioral biometrics, machine learning for authentication, or adjacent agent and security topics [3–6,9–11,13,16,25,31,32,34–39,41,56,58,59,62,68,69,74,77]. These 27 are retained in the corpus, not as sources of original performance metrics, but because their own systematic synthesis methodology (explicit search strategy, screening criteria, and structured comparison of primary literature) constitutes a methodologically grounded contribution in its own right, directly relevant to this review’s research questions on taxonomy, gaps, and prior coverage. Accordingly, these 27 works are excluded from any pooling or comparison of quantitative authentication metrics (FAR, FRR, EER, accuracy) in the Results (Section 4), where only directly reported experimental findings from the 53 primary empirical studies are synthesized; the 27 reviews are used exclusively for taxonomy construction, methodological precedent, and the related-work comparison

in Table 1. We note this composition explicitly here, rather than describing the corpus as uniformly composed of primary empirical studies, because a substantial proportion (27 of 80, that is, roughly one third) being themselves reviews is a property of this research area, in which systematic and narrative reviews are frequently cited as authoritative sources on taxonomy and terminology even within other reviews, rather than an inconsistency to be minimized. A related consequence of this composition is visible in the quality assessment (next section): because review-type studies are scored as 1 (rather than 0 or 2) on the empirical/system-grounding criterion by construction, and saturate the clarity-of-objective and referencing criteria, their total quality score is mechanically bounded to the 6–8 range, placing all 27 review-type studies in the Moderate tier regardless of their individual rigor. This is a property of how the rubric operationalizes “empirical grounding” for review-type work, not an empirical finding that all 27 reviews are of equivalent quality; per-study scores in the Supplementary Materials allow finer-grained comparison among them on the criteria that do vary (C3, C4).

3.4. Study Quality and Risk-of-Bias Assessment

Given the heterogeneous nature of the included studies—spanning empirical system evaluations, machine learning benchmarks, user studies, and protocol proposals—a formal meta-analytic risk-of-bias instrument such as RoB 2 or AMSTAR 2 was not applicable to the full corpus, since these instruments assume a common study design (e.g., randomized trials) that does not hold across this evidence base. Instead, methodological quality was assessed through an explicit criteria-based scoring rubric, applied independently by two reviewers during the full-text screening stage and reconciled by discussion. Each candidate study was scored against five criteria, with each criterion rated **0 (not met)**, **1 (partially met)**, or **2 (fully met)**, for a maximum total score of 10:

1. **Clarity of research objective** (0–2). 2: an explicit authentication goal or research question related to behavioral biometrics or continuous authentication is stated and motivated; 1: a goal is implied but not clearly articulated; 0: no identifiable authentication-related objective.
2. **Empirical or system-level grounding** (0–2). 2: the work reports experimental results, a deployed or simulated system, or a reproducible evaluation; 1: the work describes a system or methodology but with limited or secondary empirical validation; 0: purely conceptual argumentation with no system or empirical component.
3. **Reporting of performance metrics** (0–2). 2: at least one quantitative authentication metric is reported (e.g., FAR, FRR, EER, accuracy, or an equivalent measure), with sufficient detail to interpret it; 1: a metric is reported but incompletely (e.g., without sample size or operating point); 0: no quantitative authentication metric is reported.
4. **Adequate methodological description** (0–2). 2: data collection, feature extraction, and evaluation procedures are described in sufficient detail to permit independent replication or critical appraisal; 1: methodological description is present but omits material details (e.g., dataset size, feature set, or train/test protocol); 0: methods are not described in usable detail.
5. **Appropriate referencing and contextualization** (0–2). 2: the study situates its contribution within the existing literature and explicitly acknowledges key limitations; 1: the study references related work but does not discuss its own limitations; 0: the study is not contextualized against prior work.

Criteria 1 and 2 act as gatekeeping criteria: any study scoring **0** on either criterion was excluded outright, regardless of its total score, since these represent the minimum thresholds for a primary study contribution to behavioral biometric continuous authentication (an explicit authentication-related objective and at least some empirical or system-level

grounding). For studies that score at least 1 on both gatekeeping criteria, the five scores are summed into a total quality score (range 2–10), which determines how the study is treated in the synthesis:

- **High quality (score 9–10):** retained without qualification.
- **Moderate quality (score 6–8):** retained, with a quality note recorded in the extraction template; findings from these studies are interpreted with appropriate caution in the synthesis.
- **Low quality (score 2–5):** retained, but explicitly flagged as low-confidence evidence in the extraction template; findings from these studies are cited only as supporting or illustrative evidence in the Results and Discussion, never as the sole basis for a synthesized claim.

Applying this scoring rubric across all 80 included studies yielded 32 studies rated High quality (score 9–10), 44 rated Moderate quality (score 6–8), and 4 rated Low quality (score 2–5); no candidate study scoring 0 on either gatekeeping criterion was retained in the final corpus. The 4 low-quality studies span four distinct modalities rather than clustering in a single theme: continuous authentication explainability, IoT/blockchain-based intrusion detection in federated learning, touch-dynamics mobile banking authentication, and gait-based identification for elderly users in wearable healthcare. In each case, the low quality score reflects incomplete reporting of dataset size or evaluation protocol (criteria C3–C4) rather than a thematic weakness of the review’s corpus. Consistent with the treatment described above, quantitative findings from these 4 studies are not used as standalone evidence for any synthesized claim in Section 5; where cited, they are explicitly flagged as illustrative. A complete per-study quality score, including the five individual criterion scores and the final tier assignment, is provided in the Supplementary Materials (Quality_Assessment_Scores.xlsx) to allow readers to verify how each study’s quality rating affected its treatment in the synthesis.

This approach is consistent with quality assessment practices described in prior systematic reviews on continuous authentication and behavioral biometrics [4,9,10], which similarly operationalize methodological rigor through structured, criteria-based scoring rather than standardized meta-analytic instruments when the included studies are too heterogeneous for statistical pooling. In accordance with PRISMA 2020 recommendations, illustrative examples of excluded studies are provided to clarify the eligibility decisions applied at the full-text screening stage.

One study proposing a blockchain-based decentralized authentication architecture for constrained IoT nodes was excluded because, although it addressed mobile network security and device-level access control, it did not evaluate any behavioral biometric modality and reported no continuous authentication performance metrics such as False Acceptance Rate, False Rejection Rate, and Equal Error Rate. The contribution was therefore classified as an infrastructure-level security proposal rather than a behavioral biometric continuous authentication system, placing it outside the defined inclusion scope.

A second work presenting a survey of user privacy attitudes toward biometric technologies in general was excluded at the full-text stage because its empirical component was limited to a questionnaire on perceived privacy risk and did not describe or evaluate any authentication system, adaptive mechanism, or biometric pipeline. Although the attitudinal findings were contextually relevant, the absence of any system-level or performance-oriented contribution meant the study did not satisfy the inclusion criterion requiring methodological detail sufficient for critical appraisal of an authentication approach.

A third candidate, a position paper discussing ethical and regulatory implications of continuous sensor-based monitoring on smartphones, was excluded because it did not present empirical data, experimental results, or a deployable system design. The

paper argued for policy-level recommendations without validating any authentication mechanism or reporting measurable outcomes, and was therefore treated as a non-empirical commentary rather than a primary study.

These exclusions ensured that the final corpus of 80 studies consisted exclusively of peer-reviewed works reporting empirical or system-level contributions directly relevant to behavioral biometric continuous authentication on mobile or IoT platforms.

3.5. Dataset and Selection Process

Construction of the dataset of primary studies proceeded in several clearly defined stages that aligned with PRISMA-style guidance. Systematic reviews in continuous authentication and user profiling describe an analogous pipeline that begins with database querying using Boolean combinations of domain-specific terms, followed by staged screening based on inclusion and exclusion criteria, full-text assessment, and structured data extraction [4]. In behavioral biometric scoping work focused on mobile sensors, initial searches across Web of Science Core Collection, Inspec, Applied Science & Technology Source, IEEE Xplore, produced 22,537 records that were subsequently de-duplicated, screened by title and abstract, and filtered by full-text eligibility, leaving 122 studies that fully met predefined criteria [10]. PRISMA-based face liveness detection reviews likewise report a three-step protocol consisting of research question formulation, database search, and inclusion or exclusion criteria, with results summarized using a PRISMA flow diagram that documents identification, screening, eligibility, and inclusion decisions [4,16]. In line with these precedents, the present review drew on large multidisciplinary repositories such as IEEE Xplore, ACM Digital Library, ScienceDirect, Scopus, Web of Science Core Collection, Inspec, which together have been shown to cover continuous authentication, behavioral biometrics, and intelligent personal assistants comprehensively [4,5,10].

Keyword strategies were adapted from prior systematic reviews. For continuous authentication, combinations of terms such as “continuous user authentication”, “biometric authentication”, “behavioral authentication”, “continuous authentication techniques”, and “user identity verification” were employed, with Boolean operators and search filters tuned to focus on recent work [4]. For behavioral biometrics on smartphones and tablets, search strings included device and sensor terms (for example, smartphone, tablet, accelerometer, gyroscope, magnetometer, touch) together with verbs like detect, verify, authenticate, and classify, again joined by Boolean logic and database specific wildcards [10]. IPA-oriented queries followed the model of searching for variants of “intelligent personal assistant” and related phrases in the abstract field to capture both academic and applied work on assistant architectures and applications [5]. Face liveness literature informed the construction of additional AI and security related terms, such as “face anti spoofing”, “face liveness detection”, “face presentation attacks”, “artificial intelligence”, “machine learning”, and “deep learning”, which were combined using AND/OR operators in Scopus and Web of Science [16].

After query execution, records from all databases were exported to a reference manager, where duplicate removal mirrored the EndNote-based de-duplication described in behavioral biometric scoping work [10]. Remaining entries were then imported into a screening platform analogous to Covidence for title and abstract review. Following continuous-authentication survey practice, two reviewers independently screened titles and abstracts using a standardized checklist that encoded topical relevance to continuous authentication, behavioral biometrics, intelligent personal agents, and mobile or IoT contexts, along with peer-reviewed status and availability of full text in English [4]. Inter-rater reliability checks, as advocated in behavioral biometric scoping (which reported 99.9% agreement after an initial joint screening phase), were used to calibrate screening decisions before dividing

the remaining records for independent assessment [10]. Disagreements were resolved by discussion or adjudication by a third reviewer, consistent with procedures employed in both continuous authentication and IPA systematic mappings [4,5].

Full-text screening then applied a tighter set of inclusion and exclusion rules derived from prior systematic reviews. Continuous-authentication surveys required that articles directly address continuous or near-continuous user authentication, employ biometric or contextual modalities, provide sufficient methodological description, and report performance metrics such as False Acceptance Rate, False Rejection Rate, or Equal Error Rate; studies lacking methodological rigor, non-peer-reviewed documents, and duplicates were excluded at this stage [4]. Behavioral biometric scoping work mandated that eligible studies use built-in smartphone or tablet sensors for verification, detection, or authentication of the device user, and excluded work that relied on external hardware or did not implement a user verification task [10]. IPA mappings, in turn, enforced population and intervention criteria that removed any record not treating intelligent personal assistants or not containing “personal assistant” or “personal attendant” variants in the relevant fields, and required that the main subject be IPA rather than generic conversational agents [5].

Quality assessment of the candidate pool drew on checklists used in face liveness and IPA studies. For AI-based face anti-spoofing, a three-stage screening protocol culminated in a quality analysis step that removed papers failing to present clear findings, empirical support, well stated goals, or adequate referencing [16]. IPA reviews evaluated purpose of research, contextualization, literature review and related work, and methodological robustness before accepting a study as primary [5]. Continuous-authentication surveys similarly prioritized articles with transparent study designs, explicit data analysis procedures, and demonstrable relevance to the stated research questions [4]. These criteria guided the present review in selecting 80 high-quality papers, distributed across the seven thematic areas described earlier, for detailed extraction and thematic analysis.

Data extraction followed a structured template analogous to that used in continuous authentication and behavioral biometric reviews. Extracted fields included biometric and contextual modalities (for example, keystroke dynamics, gait, touch, motion, contextual usage), system type (continuous versus point of entry), learning methods (classical machine learning or deep learning architectures), evaluation metrics and reported performance, dataset characteristics (public versus proprietary, size, sensor configuration), and, where applicable, integration with intelligent personal assistant functions or IoT infrastructures [4,10,16]. For IPA related studies, additional fields captured domains of deployment, identified usability, security, and privacy challenges, and any discussion of proactivity or personalization [5]. Quality related notes, such as limitations acknowledged by the authors and gaps in experimental design, were also recorded to support later synthesis of challenges and open issues [4,10].

To enhance completeness, backward snowballing was performed on the final set of included papers. IPA mappings explicitly recommend manual searches of reference lists and related work sections to capture relevant studies not retrieved by the initial keyword queries, while continuous-authentication surveys similarly check reference lists of key articles for additional candidates [4,5]. This procedure was adopted here: reference lists of 12 key included reviews were manually checked for additional candidates (completed in February 2026), complemented by forward citation tracking via Google Scholar for five seminal papers (completed in March 2026). Newly identified records were subjected to the same screening and quality assessment pipeline as the database-derived set. Together, these steps yielded a curated dataset of 80 studies that collectively cover continuous mobile user authentication via behavioral and multimodal biometrics, intelligent personal assistants,

and supporting machine learning and security mechanisms, assembled under a transparent, PRISMA consistent selection process [4,5,10,16].

3.6. Thematic Areas of Analysis

Organizing the evidence base into coherent themes is essential for answering the research questions defined earlier and for structuring the subsequent results. Existing systematic reviews on continuous authentication and related topics provide concrete precedents for thematically grouping primary studies after PRISMA-based selection. A comprehensive review of continuous authentication explicitly structures its synthesis into sections on supervised and unsupervised learning, physiological biometrics, behavioral biometrics, multimodal biometrics, context-aware techniques, other emerging approaches, open issues and challenges, future research directions, related surveys, and key insights, thereby demonstrating how methodological, modality-specific, and forward-looking strands can be separated yet jointly analyzed [4,16]. Work focused on touch dynamics similarly devotes dedicated sections to usability and ergonomic factors, single modality analyses for keystroke, swipe and signature, multimodality, datasets, attacks, and open problems, again illustrating that thematic organization can follow both modality and cross-cutting concern dimensions [32].

For the present review, thematic categorization builds directly on such prior structuring. Continuous-authentication surveys show that physiological, behavioral, multimodal, and context-aware techniques form natural first level themes, each associated with distinct sensing infrastructures, feature representations, and classifier choices [4,9]. Within behavioral biometrics, further subdivision into keystroke dynamics, touch or swipe interaction, and motion- or gait-based schemes is supported by survey level analyses that treat these modalities separately when discussing data acquisition, feature extraction, classification, and usability factors [4,32]. At the same time, context-aware authentication, intelligent personal assistant architectures, privacy-preserving and blockchain-based mechanisms, and emerging AI-driven defenses against spoofing are consistently treated as separate sections or research questions in existing surveys, which motivates grouping them into higher level themes that cut across individual biometric traits [4,5,16].

Systematic review methodology texts emphasize that thematic areas should be derived after data extraction, using techniques such as thematic analysis and narrative synthesis to identify common patterns and gaps [4,16]. In continuous authentication, extracted fields such as biometric modality, learning method, performance metrics, dataset type, and reported challenges naturally support a thematic map in which one axis corresponds to modality classes and another to system level issues like usability, security, privacy, and deployment constraints. A PRISMA-based face anti-spoofing review, for instance, first classifies primary works by attack type and method category, then synthesizes trends within each group (for example, handcrafted texture methods, deep learning-based methods, and multi-cue methods), before discussing open problems and future research for each cluster [16]. A similar logic informs the current thematic framework, where seven overarching areas are defined to capture: (i) classical physiological and knowledge or possession-based factors; (ii) core behavioral modalities on mobile devices; (iii) context-aware and multimodal fusion schemes; (iv) intelligent personal assistants and agent architectures involved in authentication; (v) machine learning techniques from classical algorithms to deep and emerging approaches; (vi) security threats, error metrics, and evaluation issues; and (vii) deployment in wearables, IoT, and blockchain enabled environments. Evidence from prior surveys confirms that such a decomposition aligns with how research communities already segment their discussions and research questions [4,5,9,16].

Methodological guidance from continuous authentication and behavioral biometric scoping reviews also shapes how studies are distributed across themes. After inter-rater screened inclusion, one review explicitly codes each paper for modality, system type (continuous or static), and research objective, then aggregates counts to identify which areas are well or under represented [4,10]. In touch dynamics, a similar coding process distinguishes works centered on usability evaluation, ergonomic impact, security vulnerabilities, and dataset provision, and then associates each primary study with one or more of these categories [32]. Building on these practices, the present review codes each of the 80 included papers against every one of the seven thematic areas it substantively addresses, rather than restricting assignment to a single dominant theme. A study that combines motion-based authentication with an explicit adversarial-robustness analysis, for example, is tagged under both the behavioral-modality theme and the security-threats theme. The final distribution table therefore reports theme-level counts that sum to more than 80, reflecting genuine multi-theme coverage in the underlying literature rather than a coding artefact, mirroring the way earlier reviews tabulate inclusion by modality or method [4,10,32].

Several surveyed articles explicitly recommend connecting methodological and technical strands within the thematic structure. Continuous authentication work that spans supervised and unsupervised learning, modality-specific analysis, context-aware schemes, and emerging methods argues that future surveys should relate algorithmic choices directly to biometric categories and application domains [4]. AI-based face anti-spoofing reviews further suggest that emerging topics such as domain generalization, federated learning, and explainable AI deserve separate yet linked thematic treatment because they cut across specific biometric modalities [16]. Intelligent personal assistant mappings define their own general and specific research questions around state of the art, taxonomies, usages, privacy and security issues, and architectures, thereby providing an additional thematic lens centered on assistant ecosystems rather than biometrics alone [5]. Incorporating these recommendations, the thematic areas adopted here are explicitly constructed to accommodate both modality grounded discussion and higher level issues like intelligent agents, privacy, and adversarial robustness.

Finally, the planned analyses within each theme follow synthesis techniques already validated in related reviews. Continuous-authentication surveys employ narrative synthesis and, where applicable, meta analysis of error rates across modalities, while face anti-spoofing work uses comparative tables to contrast methods, datasets, and performance within each attack class [4,16]. Behavioral biometric scoping reviews present modality-specific sections that link usability metrics, ergonomic factors, and security vulnerabilities to classification results, culminating in discussion of open research problems [32]. By adopting similar within theme analyses and cross-theme comparisons, the current review aims to maintain methodological continuity with existing systematic studies while offering an integrated view across continuous mobile authentication, behavioral biometrics, and intelligent personal agents [4,5,10,16].

4. Results

4.1. Thematic Distribution

Table 6 summarizes the distribution of included studies across seven thematic areas identified through systematic analysis of the primary literature. IoT, wearables, blockchain and agent-based systems represent the largest research cluster, encompassing work on continuous authentication across smart homes, healthcare IoT, vehicular networks, and distributed learning paradigms (32 studies). Multimodal fusion strategies and machine learning methods form the second major group (18 studies), reflecting the widespread adoption of deep learning architectures such as LSTMs, CNNs, and transformers for behav-

ioral biometric processing. Security threats and defense mechanisms constitute another significant theme (17 studies), addressing adversarial robustness, spoofing attacks, and privacy preservation. Within behavioral biometrics, keystroke dynamics (12 studies) and gait and motion sensor-based approaches (11 studies) dominate, followed by touch and swipe gesture recognition (five studies). Classical authentication methods based on passwords and PINs receive minimal attention (one study), consistent with the review’s focus on behavioral and continuous approaches. Note that some studies address multiple themes and are counted accordingly: the seven theme-level counts sum to 96 rather than to the 80 included studies, indicating 16 additional theme-tags contributed by studies that substantively span two or more thematic areas. For example, a motion-based continuous authentication scheme that explicitly analyzes adversarial robustness [14] is tagged under both the behavioral-modality and the security-threats themes, while a mobile-money authentication design combining a SIM-based possession factor with touch and accelerometer signals [2] is tagged under both the classical-authentication and the behavioral-modality themes. This reflects the interdisciplinary nature of continuous authentication research rather than double-counting error.

Table 6. Thematic distribution of the 80 included studies across seven research areas.

Thematic Area	Studies	Percentage
IoT, Wearables, Blockchain, Intelligent Agents	32	33.3%
Multimodal Fusion and Machine Learning Methods	18	18.8%
Security Threats, Attacks, and Defenses	17	17.7%
Keystroke Dynamics	12	12.5%
Gait and Motion Sensor-Based Biometrics	11	11.5%
Touch and Swipe Gesture Biometrics	5	5.2%
Classical Authentication (Password, PIN, Tokens)	1	1.0%
Total (with overlap)	96	100.0%

A more granular breakdown of the included studies by biometric modality and machine learning method is provided in Tables 7 and 8. Among biometric modalities, gait (13 studies) and multimodal approaches (11 studies) are most represented within the behavioral biometrics category, while survey and IoT security papers constitute the largest non-biometric clusters. Regarding algorithmic methods, CNN-based architectures (11 studies) and classical ML approaches (20 studies) dominate, with federated learning emerging as a growing paradigm (4 studies).

Figure 3 illustrates the thematic overlap among the three principal research clusters identified in the reviewed literature. Behavioral biometrics represents the largest domain with 45 studies, of which 16 address this theme exclusively. The intersection of behavioral biometrics and continuous authentication encompasses 17 studies, reflecting the close relationship between these two fields. The central region where all three domains converge comprises 8 studies—the most integrative works that simultaneously address behavioral biometrics, continuous authentication, and IoT or intelligent agent architectures. The relatively small size of the continuous authentication-only cluster (three studies) confirms that this topic is rarely studied in isolation from behavioral or IoT contexts.

Thematic Overlap in Reviewed Studies

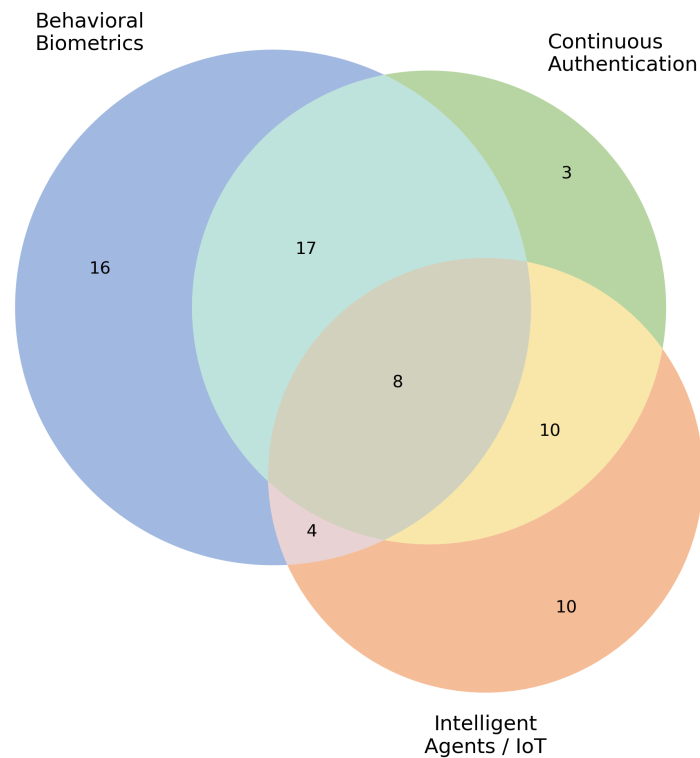


Figure 3. Thematic overlap of reviewed studies across three principal research domains. Numbers indicate the count of studies per region; the seven regions shown sum to 68 of the 80 included studies, since the remaining 12 studies address topics outside all three of these specific domains (e.g., purely classical knowledge-based authentication or general-purpose machine learning methods not tied to a specific modality covered here) and are therefore not represented in this diagram.

Table 7. Distribution of included studies by biometric modality.

Biometric Modality	Studies	Percentage
Survey	24	25.0%
IoT Security	13	13.5%
Gait	13	13.5%
Multimodal	11	11.5%
Agent	10	10.4%
Keystroke	8	8.3%
Face	5	5.2%
ML Security	5	5.2%
Touch	4	4.2%
Voice	1	1.0%
EEG	1	1.0%
Total	95	100%

Note: Survey = systematic review papers; IoT Security = IoT/blockchain/V2X; Agent = intelligent personal assistant studies. As with Table 6, several studies combine more than one modality (e.g., a multimodal scheme fusing keystroke and motion signals) and are therefore counted once under each relevant modality category; this is why the total exceeds the 80 included studies and should not be read as double-counting error. Note also that the 27 review papers retained as primary studies (Section 3.2) do not map one-to-one onto the “Survey” category above: 24 are coded as Survey, while 3 are coded under domain-specific categories (Agent or IoT Security) because their content is predominantly a systematic mapping of intelligent personal assistant or IoT/blockchain literature respectively, rather than a general continuous-authentication survey.

Table 8. Distribution of included studies by machine learning method.

ML/Algorithm Method	Studies	Percentage
Survey	23	24.0%
Classical ML	20	20.8%
No ML	18	18.8%
CNN	11	11.5%
Deep Learning	8	8.3%
Federated Learning	4	4.2%
Random Forest	3	3.1%
LSTM	3	3.1%
Transformer	2	2.1%
SVM	2	2.1%
Total	94	100%

Note: No ML = theoretical/protocol/blockchain papers; Survey = review papers without specific model. As with Tables 6 and 7, studies employing more than one learning method (e.g., a hybrid RNN+SVM scheme) are counted under each method used, which is why the total exceeds the 80 included studies.

4.2. Publication Trends

The systematic search strategy and subsequent screening process produced a focused corpus of studies that enables a structured analysis of publication activity, selection decisions, and geographical coverage. The following sections summarize the temporal distribution of publications, the article selection pipeline, and the geographical distribution of the included studies.

Figure 4 presents a bar chart of the number of publications per year retrieved from the configured databases, primarily between 2017 and 2025, supplemented by a single earlier paper recovered through backward snowballing (Section 3.5) that falls outside this window. The yearly counts show a pronounced peak in 2020 (18 of the 80 included studies), followed by 16 studies in 2021, 15 in 2022, 14 in 2023, 6 in 2024, 4 in 2025, 5 in 2019, 1 in 2017, and 1 in 2015 (via snowballing). We report this distribution descriptively and avoid attributing it to a single cause: publication counts in any given year reflect a mix of factors that this review cannot disentangle, including genuine shifts in research activity, editorial and review timelines at target venues, and indexing lag for the most recent years. In particular, the lower counts for 2024 and 2025 are at least partly attributable to indexing lag, since our search concluded in April 2026 and the most recent publications may not yet be fully indexed in the queried databases; we therefore do not interpret this decline as evidence that research interest in the field is waning.

To further characterize the thematic focus of the reviewed literature, Figure 5 presents a word cloud generated from the titles of all 80 included studies. Dominant terms such as *Authentication*, *Mobile*, *Security*, *Behavioral Biometrics*, and *Deep Learning* confirm that the corpus is tightly centered on continuous and behavioral approaches to mobile security, with emerging prominence of *Gait*, *Keystroke*, *Federated Learning*, and *Blockchain* reflecting recent research directions.

Figure 6 depicts the PRISMA-style flow diagram summarizing the article selection process from identification to final inclusion. In the identification phase, the five search strings described in Section 3.2 returned approximately 1744 records across the six queried databases (Scopus, Web of Science, IEEE Xplore, ACM Digital Library, ScienceDirect, and SpringerLink) [4,5,10,16]. After removing approximately 1364 duplicate records using Zotero with manual verification, 380 unique records remained for title and abstract screening. Screening, conducted independently by two reviewers, excluded 70 records, leaving 310 articles for full-text eligibility assessment. Full-text assessment excluded 230 articles: 100 for insufficient methodological rigor, 30 for not being peer-reviewed, and 100 identi-

fied as duplicates not caught at the initial deduplication stage. This yielded a final set of 80 included studies [4,10,16,39].

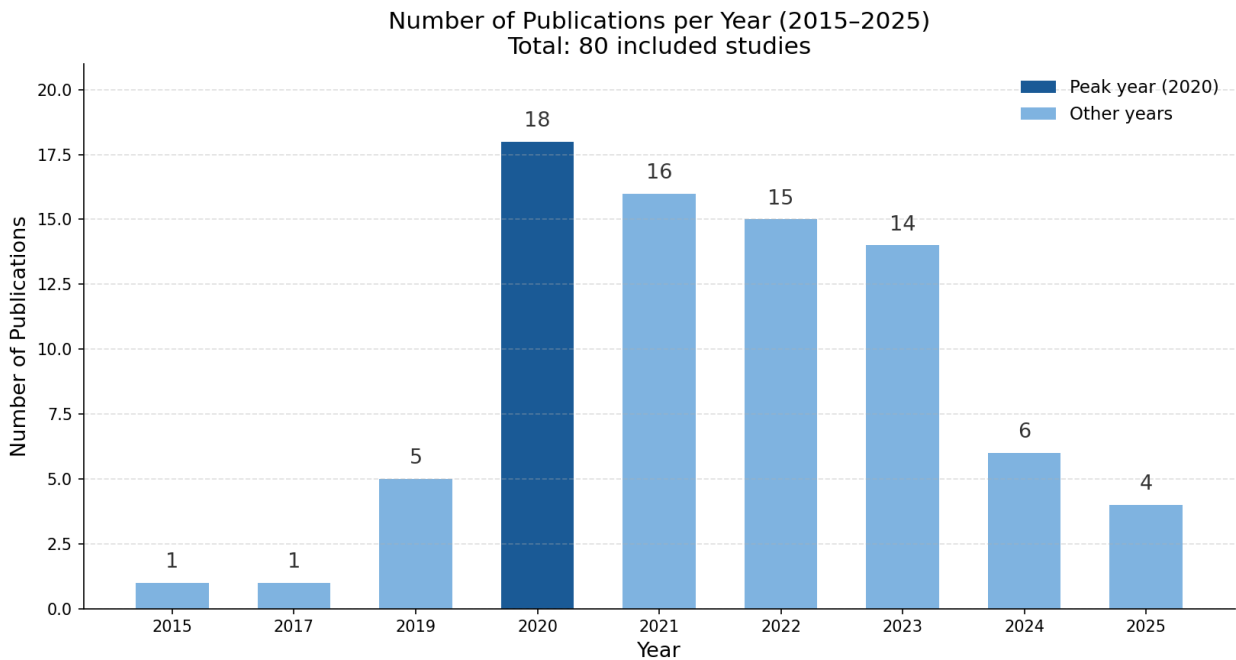


Figure 4. Number of publications per year (2015–2025).

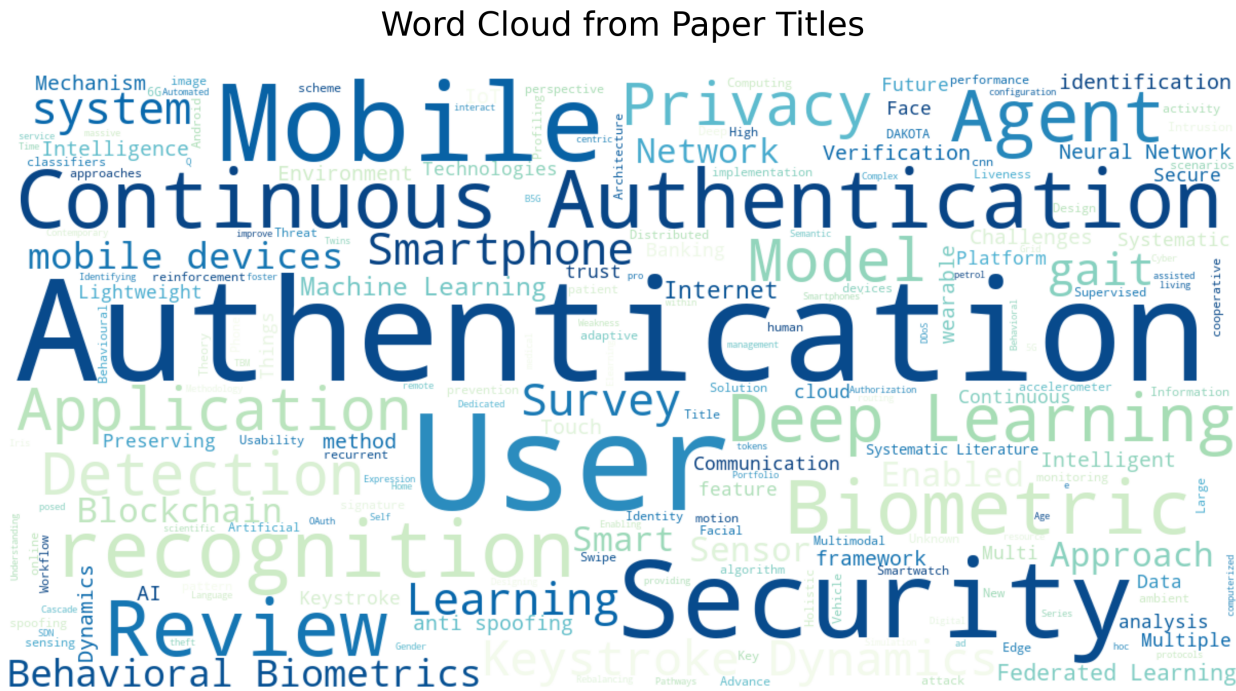


Figure 5. Word cloud generated from titles of the 80 included studies, illustrating dominant research themes.

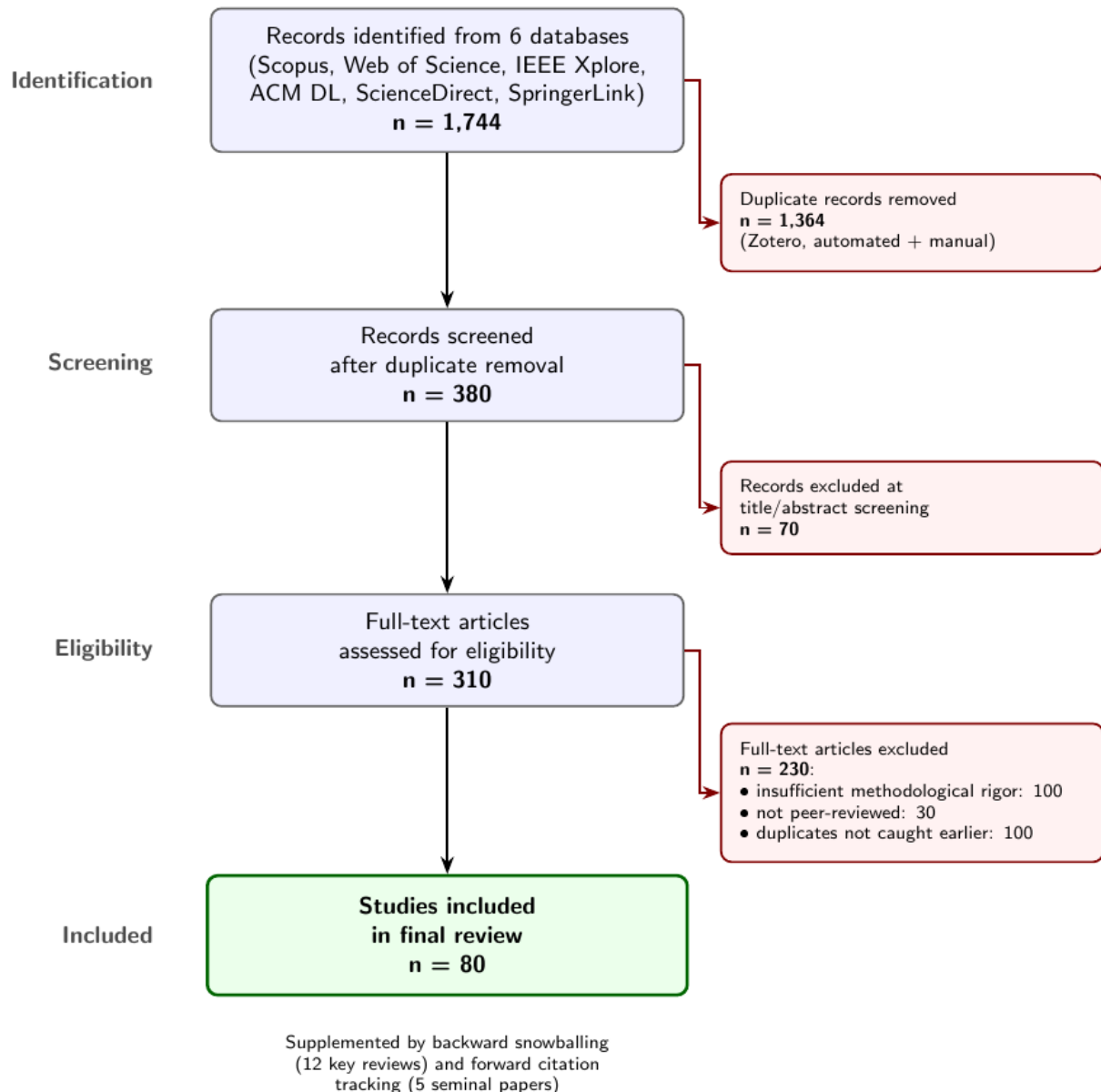


Figure 6. PRISMA 2020 flow diagram of the article selection process.

4.3. Geographical Distribution

Figure 7 summarizes the geographical distribution of the 80 included publications. China contributes the largest share with 13 publications, followed by the United States with 9 and India with 7. South Korea and Spain each account for 6 studies, while Australia contributes 5. The UAE, Saudi Arabia, the United Kingdom, Turkey, and Greece each contribute 3 publications, followed by Israel, Brazil, Egypt, and Thailand with 2 each [3,4,10]. This distribution reflects the global nature of research on mobile authentication and behavioral biometrics, with notable activity across Asia, Europe, and Oceania [3–5,32].

Taken together, Figures 4–7 provide a compact overview of how the literature has grown over time, how the final study set was derived using a transparent PRISMA-based workflow, and how research efforts are distributed geographically.

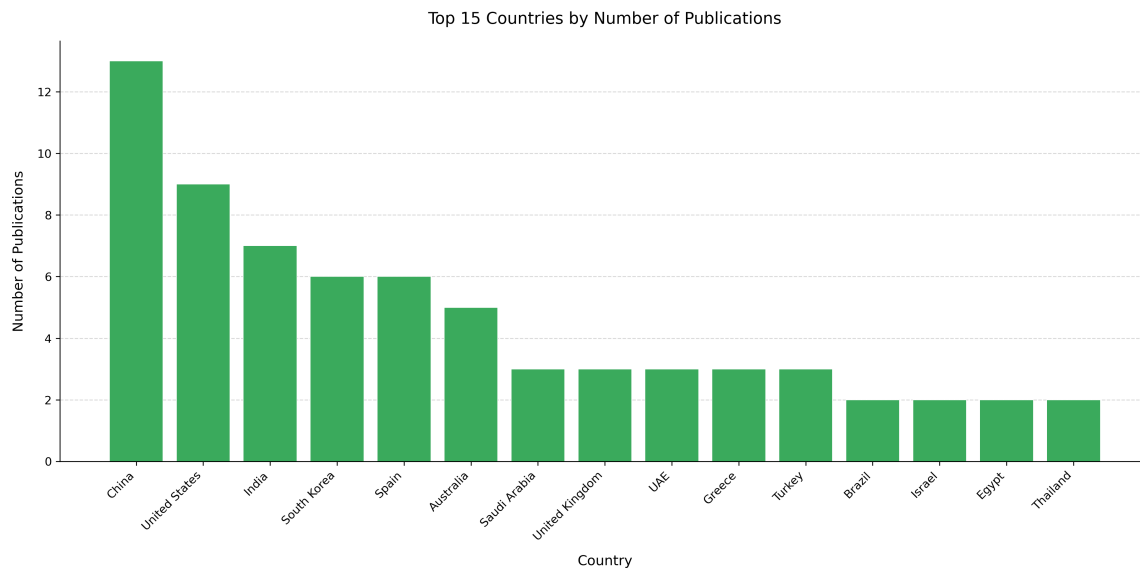


Figure 7. Geographical distribution of included studies by country of corresponding author (top 15 countries).

5. Discussion

5.1. Key Findings and Their Interpretation

The reviewed literature converges on a structural paradox that no current system has fully resolved: continuous authentication becomes more accurate as more behavioral data are collected, yet the richness of those data simultaneously increases privacy exposure, energy consumption, and attack surface. This is not a design trade-off amenable to straightforward optimization—it is structural.

The temporal density, contextual richness, and multimodal coverage that make behavioral biometrics effective for security are precisely the properties that render compromised profiles most damaging. Unlike passwords, behavioral templates encode daily routines, demographic attributes, and activity patterns that cannot be reset once leaked [4,9].

The multimodal advantage documented across the reviewed studies is real but systematically fragile. Under benign evaluation conditions, fusion of keystroke, touch, motion, and contextual signals consistently reduces error rates—LSTM Siamese architectures report TAR values near 99.98% at 0.1% FAR across eight modalities [37].

However, this advantage collapses under adversarial conditions. Generative attacks drive touch dynamics EER from 5% to 50%, and keystroke dynamics EER from 28% to 84% [9]. Treadmill-assisted gait imitation inflates FAR from 5.8% to 43.66%, and ECG template mapping achieves attack success rates above 90% [4,9].

The implication is uncomfortable: performance figures reported in the literature are measured almost exclusively under benign protocols, meaning the reported state of the art is likely considerably weaker in real adversarial deployment than published EER values suggest.

Adopting deeper learning architectures improves accuracy but shifts rather than eliminates the threat surface. Transformer and LSTM models outperform classical classifiers on large datasets and support end-to-end representation learning from raw sensor streams [37]. Yet these architectures introduce model-level vulnerabilities—membership inference, reconstruction, and model extraction—that are absent or substantially less severe in simpler kNN and SVM pipelines [9,36]. This creates a second paradox: the models that best discriminate genuine users from impostors are also the models that most readily expose user-specific information when queried by an adversary. Federated learning reduces raw data exposure by keeping sensor streams local, but introduces poisoning and Byzantine attack risks at the

aggregation layer [69]. The reviewed literature has not produced a satisfying resolution to either paradox, and the tendency to report accuracy gains without reporting corresponding privacy or adversarial costs represents a methodological gap that future work must address.

The integration of intelligent personal agents with continuous authentication is the least mature theme in the reviewed corpus. Despite IPA appearing in the title of this review, only 8 of the 80 included studies simultaneously address behavioral biometrics, continuous authentication, and agent or IoT architectures—as illustrated by the small central region of Figure 3. No study was identified that evaluates a deployed commercial IPA system such as Apple Siri, Google Assistant, or Amazon Alexa as an active component of a continuous authentication pipeline with reported FAR or FRR metrics. This absence is significant because IPAs represent the most natural ambient deployment context for continuous behavioral monitoring: they already coordinate sensors, maintain interaction histories, and mediate cloud communication. The failure of the literature to engage with this context empirically means that the security and privacy implications of IPA-mediated authentication remain speculative [4,5].

Privacy-preserving mechanisms are proposed throughout the reviewed literature but are almost never evaluated end-to-end within a full authentication pipeline. Bloom filter encodings, cancelable biometrics, and zero-knowledge proof protocols appear as recommended safeguards in approximately fifteen of the reviewed papers, yet in nearly every case they are discussed in isolation—either as theoretical constructs or as components tested outside the behavioral authentication context. Studies proposing ZKP-based protocols do not evaluate them against the keystroke or gait modalities reviewed here; studies evaluating touch dynamics or gait authentication do not implement cancelable templates. The consequence is that the privacy-utility trade-offs of these mechanisms under realistic continuous authentication workloads—with adversarial inputs, temporal drift, and energy constraints—remain empirically unknown [4,9].

Real-world deployment introduces a further set of challenges that the reviewed literature addresses only indirectly. Almost all reported performance figures come from controlled data-collection sessions with cohorts of well under 100 participants (Section 4), rather than from production deployments spanning heterogeneous devices, operating-system versions, and years of behavioral drift. Practical deployment must additionally contend with device heterogeneity (differing sensor sampling rates and touchscreen digitizers across manufacturers), intermittent connectivity that affects cloud-assisted or hybrid architectures (Section 2.5.2), and the operational cost of re-enrollment when behavioral drift degrades accuracy over months of use. None of the 80 studies reviewed here reports a multi-year, multi-device field deployment; the closest approximations are the larger, unconstrained-condition datasets discussed in Section 2.4.3 [47], which improve ecological realism but still represent single-session or short-longitudinal data collection rather than sustained production use.

Generative AI introduces threats that go beyond the generative adversarial attacks already discussed above. The EEG-synthesis and face-spoofing GANs surveyed in Section 2.6.3 demonstrate that generative models can already produce convincing synthetic biometric signals for physiological modalities [16,68]; the same generative capability, applied to behavioral time-series (keystroke timing sequences, touch-gesture trajectories, accelerometer-based gait signals), would in principle allow an attacker to synthesize plausible impostor behavioral traces rather than only replaying or imitating recorded ones. None of the 80 studies reviewed here evaluates a behavioral-modality authentication system against a generative behavioral-trace attack of this kind; the closest evidence is the touch- and keystroke-dynamics work showing that classifier decision boundaries can be manipulated by crafted adversarial inputs [9], which is a related but distinct threat model

from a generative model that produces an entire synthetic behavioral session. This gap is consequential precisely because large generative models capable of producing such synthetic sequences have become substantially more accessible since most of the reviewed studies were conducted, suggesting that adversarial-robustness evaluation protocols for behavioral biometrics will need to be revisited as this capability matures.

A recurring obstacle to closing the gaps above is the absence of shared benchmark datasets. The studies reviewed in Sections 2.4.1–2.4.4 and Table 3 draw on a fragmented set of study-specific corpora (from small in-house collections of fewer than 50 participants to larger resources such as OU-ISIR and whuGait for gait, and TypeNet’s 26,000-participant free-text keystroke corpus [44]), rarely reused across research groups and almost never annotated with a common adversarial or generative-attack protocol. This fragmentation has three consequences for the field. First, reported EER and accuracy figures are not directly comparable across studies, since dataset size, session count, and environmental control vary substantially (Section 2.7.1). Second, no dataset reviewed here jointly provides genuine sessions, imitation-attack sessions, and generative-attack sessions for the same modality, which is precisely the resource that would be needed to evaluate the generative-threat gap identified above. Third, the larger, more ecologically valid datasets (e.g., unconstrained-condition gait corpora [47]) tend to be modality-specific, so no dataset in the reviewed corpus supports the kind of multimodal, longitudinal, adversarially annotated benchmark that would let the community compare architectures, fusion strategies, and privacy-preserving mechanisms on a common footing.

Taken together, these observations point to a single underlying pattern rather than three unrelated findings. The reviewed literature predominantly treats behavioral modalities, machine learning architecture choices, agent and IoT integration, and privacy-preserving mechanisms as separable research domains, each evaluated against its own narrow set of metrics (accuracy or EER for modalities, FAR/FRR for fused systems, formal privacy guarantees for cryptographic protocols), rather than as interacting components of a single deployed system. The deep learning paradox (higher accuracy traded against greater exposure to membership inference and reconstruction), the IPA integration gap (only 8 of 80 studies jointly addressing behavioral biometrics, continuous authentication, and agent or IoT architectures, Figure 3), and the isolated evaluation of privacy-preserving mechanisms are each, individually, a gap within one domain; considered jointly, they indicate that the domains most likely to compound each other’s weaknesses in a real deployment, accurate-but-leaky models, agent-mediated data flows, and unvalidated privacy safeguards, are precisely the domains the literature has not yet studied together. This is consistent with the architectural analysis in Section 2.5.2, where the TEE boundary protects the model and decision logic but not the feature-extraction and memory-leakage surfaces that connect the device agent to edge and cloud tiers: the unresolved overlap between domains in the literature mirrors an unresolved overlap between attack surfaces in the architecture itself.

5.2. Research Gaps Identified

Persistent shortcomings documented across continuous authentication, behavioral biometrics, and intelligent personal assistants expose several concrete gaps that future work must address. Continuous User Biometric Authentication Systems are repeatedly described as facing unresolved trade-offs between usability and security, with FAR and FRR still difficult to balance in ways that satisfy both security requirements and user comfort [4]. Comprehensive strategies for usability management across modalities, including systematic evaluation protocols and design patterns that incorporate human intervention smoothly into continuous workflows, remain largely absent, despite explicit calls for more holistic approaches [4,32]. In parallel, IPA research reports a lack of standardized measures for

comparing assistant applications, which complicates assessment of usability, security, and privacy properties and makes empirical comparison of different architectures mostly ad hoc [5].

Security analyses pinpoint a second cluster of open issues. Behavioral biometrics such as touch dynamics and keystroke dynamics have been shown to be vulnerable to generative and adversarial attacks that drastically elevate Equal Error Rates, yet there is little evidence of robust, systematically evaluated defense mechanisms tailored to these attack models [4]. Machine learning-based continuous authentication systems, particularly those using SVM and kNN, are further exposed to sample inference, reconstruction, membership inference, and model extraction attacks, but protection strategies at the model and API level are still only sketched conceptually rather than rigorously integrated into deployment frameworks [9]. For face liveness detection, current deep learning-based techniques do not generalize well to invisible or synthetic attacks, and calls for domain adaptation, domain generalization, and combined handcrafted plus deep feature strategies underscore that reliable cross-domain robustness remains an open research area [6,16,77].

Privacy and ethical dimensions introduce another set of gaps. Continuous-authentication surveys stress that persistent collection and storage of biometric and contextual data may infringe individual privacy rights, yet concrete, deployable privacy-preserving architectures for CA remain underdeveloped [4]. Practical work on keystroke dynamics has proposed feature-scoring methods that limit exposure of raw user data by relying only on aggregate statistics, for example, requiring only mean values from other users, which appears to offer a modest path toward reducing information leakage in continuous pipelines [78]. Nevertheless, such methods have been evaluated on relatively small mobile datasets and it is not yet clear how they stand up to reconstruction or membership inference attacks in long-term, real-world deployments. Work on mobile phone authentication underscores that behavioral authentication becomes especially problematic when algorithms or data are outsourced to third parties, and warns that educating users alone is insufficient if underlying systems do not incorporate unlinkability, non-invertibility, and revocability of stored information [2]. Although future directions propose zero-knowledge proof, quantum cryptography, and other cryptographic primitives as promising building blocks, detailed designs and evaluations of such schemes in mobile and assistant-mediated continuous authentication settings are still lacking [4].

From a modality and dataset perspective, several deficiencies stand out. Studies on gait recognition note that manual feature extraction is time consuming, requires substantial domain expertise, and lacks standardization, while deep learning-based approaches, although more accurate, tend to have high model complexity unsuited to wearable devices with limited computational resources. Existing gait datasets such as OU ISIR and whuGait provide important baselines but do not cover very large populations, long-term temporal variability, or diverse environmental conditions, leaving scalability and temporal robustness insufficiently explored [44]. In face presentation attack detection, publicly available datasets remain limited in volume and diversity, cover only a subset of known spoofing types, and rarely capture real-world acquisition variability, which constrains the realism of evaluation and hinders progress on unseen attack robustness [16,77].

Methodological gaps at the review and system design levels also emerge. Continuous-authentication surveys themselves acknowledge that research has focused heavily on supervised learning and standard metrics such as FAR and FRR, with little comparative analysis across physiological, behavioral, and multimodal categories and few additional measures that capture practical success factors like scalability and long-term usability [4]. Behavioral biometric scoping work on mobile sensors documents the scarcity of high quality, ecologically valid studies that use built-in smartphone or tablet sensors for real-

world verification, detection, or authentication, despite an initial pool of over twenty thousand records [10]. IPA mappings similarly report an absence of measures enabling direct comparison of assistant applications, limited attention to specific user profiles such as the elderly, and a dearth of systematic evaluations of domains like education, health, and privacy, all of which represent fertile ground for integrating continuous behavioral authentication but remain under investigated [5].

Energy and deployment constraints constitute a further unresolved domain. Evidence on smartphone battery usage shows that numerous sensors, including light, proximity, gyroscope, and digital compass, consume substantial power, and that sensory data collection itself increases electricity usage, yet only a few studies provide thorough power analyses for CUBAs despite their requirement for sensors and processing to remain active throughout full sessions [4]. As mobile devices evolve into general purpose smart platforms and MPAS designs grow more complex, research gaps around systematic energy measurement, energy aware modality selection, and scalable deployment of continuous authentication in resource-constrained environments become increasingly salient [2,4].

Finally, cross-domain integration of continuous authentication with intelligent personal assistants and IoT or wearable ecosystems is still in its infancy. IPA reviews call for future work on usability, security, and privacy challenges, on proposing tools and frameworks to facilitate new assistant skills, and on exploring under-served domains such as health, education, and ambient assisted living, yet they do not systematically connect these agendas to biometric continuous authentication [5]. Continuous-authentication surveys, for their part, identify multimodal, context-aware, and blockchain-integrated schemes as promising directions, but provide limited guidance on how to embed these techniques into agent-based architectures that span smartphones, wearables, fog or cloud infrastructures, and assistant platforms [4,6]. Bridging these gaps will require coordinated efforts that jointly address modality robustness, privacy-preserving protocols, energy constraints, and assistant-centric interaction models within a unified continuous authentication framework.

Three gaps stand out as the most critical barriers to responsible large-scale deployment of behavioral biometric continuous authentication. First, no published study evaluates a complete pipeline—from raw sensor acquisition through feature extraction and template protection to a final authentication decision—under both adversarial and privacy threat models simultaneously. Existing work partitions these concerns into separate papers, leaving their interactions unmeasured. Second, no standardized benchmark dataset exists that combines realistic attack simulations, longitudinal enrollment scenarios, and explicit energy and latency measurements, making cross-study comparison of security and usability claims unreliable [4,10]. Third, and most fundamentally, user acceptance of privacy-preserving mechanisms has not been studied. Research confirms that privacy concerns and perceived response costs reduce intention to adopt BBCA [1], yet no study has tested whether the deployment of cancelable biometrics, ZKP protocols, or federated architectures actually increases user willingness to engage with continuous authentication in practice. Closing these three gaps—integrated pipeline evaluation, standardized adversarial benchmarks, and user-centered privacy studies—represents the most consequential research agenda for the field before BBCA can be deployed responsibly at scale [1,4,10,79].

These substantive gaps should be read alongside two review-level limitations. First, as noted in the Methodology, no quantitative publication-bias correction was applied; reported EER and accuracy figures synthesized above are drawn from the published literature and may skew optimistic relative to unpublished negative or null results, including proprietary industry deployments that are not submitted for peer review. Second, quality scoring and thematic coding, while performed independently by two reviewers with disagreements resolved by discussion, retain an irreducible element of subjective judgment

inherent to narrative synthesis across heterogeneous study designs; a different reviewer team applying the same rubric could plausibly shift individual studies between adjacent quality tiers, though the overall High/Moderate/Low distribution (32/44/4) is unlikely to change substantially.

6. Ethical Considerations

6.1. Confidentiality of Behavioral Data

Continuous-authentication pipelines depend on dense streams of behavioral and contextual signals, and these streams inherently encode sensitive information about individuals and their activities. Continuous schemes based on gait recognition, stylometry, touch dynamics, and related modalities operate on daily-routine data and current actions, which cannot be permanently changed if compromised, unlike passwords that users can reset.

Keystroke dynamics in particular can leak demographic attributes such as age, gender, and handedness, exposing identity and contextual privacy when templates or models are accessed by adversaries [9]. Context-aware modes extend this exposure by logging location, browsing history, application usage, and network interactions, so behavioral profiles jointly assembled from biometric and contextual streams may reveal long-term patterns of movement and device usage that users did not intend to disclose [4,9].

IoMT and eHealth settings add another layer of sensitivity, since remote patient monitoring architectures collect physiological and medical data and transmit them over wireless links that are vulnerable to eavesdropping, denial of service, and insider misuse, while many systems still rely on centralized servers that represent single points of failure and attractive targets for data theft [8,79].

Model and system design decisions further shape confidentiality risks. In federated learning for IoMT, devices locally train models and periodically upload weights and gradients to an aggregation server; although this paradigm is intended to preserve device-side privacy, untrusted servers can exploit received parameters to reconstruct private local data, especially when adversarial techniques perturb initial parameters or labels over time to disclose sensitive information [8]. Continuous authentication deployments that use classifiers such as Support Vector Machines and k Nearest Neighbors in cloud settings can store actual user samples within authentication profiles, enabling sample inference and reconstruction attacks that re-identify individuals from reconstructed biometric traces [9]. Blockchain-based smart home and eHealth frameworks illustrate complementary confidentiality issues: while public ledgers provide tamper resistant storage of authentication and medical records, transaction data remain globally visible, so deanonymization attacks that link on-chain records with external databases can reveal user identities and associated activities, and loss of private keys can permanently deprive users of control over their own data [60,79].

Several technical mechanisms have been proposed to protect behavioral data within continuous authentication ecosystems. Bloom filters offer a space efficient, stochastic data structure for membership queries with controlled false positives and constant time access; recent work applies Bloom filters to biometric templates, including adaptive constructions that create irreversible and unlinkable representations, thereby reducing the risk that leaked templates can be inverted to recover original behavioral signals. Cancelable biometrics provide template non-reversibility and biometric salting by transforming raw biometric data into distorted domains through operations such as random projection, fingerprint mixing, or BioHashing-based mappings, so that compromised templates can be revoked and reissued without exposing the underlying behavioral pattern. Zero-knowledge-proof-based authentication protocols allow one party to prove legitimacy to another without disclosing secrets, cryptographic keys, or explicit authentication information; in continuous or

cloud-assisted settings this prevents servers, including curious or malicious insiders, from learning user templates or metadata, even if they participate in the verification process [4]. Table 9 provides a structured comparison of the principal template protection methods identified in the reviewed literature, evaluating each mechanism across key privacy and operational dimensions. As with Tables 1 and 2 above, the Irrev./Unlink./Revoc. and Overhead columns are qualitative, ordinal judgments drawn from the cited literature rather than measured benchmarks on a shared implementation; the “(estimated)” and “N/A” entries for ZKP and Federated Learning respectively reflect that no single reviewed study reports these mechanisms’ overhead on a directly comparable workload, so cross-row comparisons in this table should be read as relative orderings rather than absolute cost figures.

Table 9. Comparison of biometric template protection methods.

Method	Core Principle	Irrev.	Unlink.	Revoc.	Overhead	Best Suited For
Cancelable Biometrics	Transform raw biometric via non-invertible function (BioHashing, random projection)	High	High	Yes	Low	Keystroke, fingerprint templates
Bloom Filter Encoding	Encode biometric into probabilistic bit-array; constant-time queries	High	High	Yes	Low	Binary biometric features, iris codes
Zero-Knowledge Proof (ZKP)	Prove identity without revealing secret or template to server	High	High	Yes	High	Medium–High (estimated)
Federated Learning	Train model locally; share only gradients/weights, not raw data	Medium	Medium	Partial	Medium	N/A
Fuzzy Commitment	Bind cryptographic key to biometric using error-correcting codes	High	Medium	Yes	Medium	Cryptographic key binding
Differential Privacy	Add calibrated noise to stored data or model updates	Medium	High	Partial	Low	Blockchain, federated analytics
Quantum Cryptography	Use quantum key distribution for theoretically unbreakable encryption	Very High	Very High	Yes	Very High	Theoretical/future

Note: ZKP—zero-knowledge proof. Overhead refers to computational and communication cost relative to unprotected baseline, synthesized qualitatively from the reviewed literature on these mechanisms [4,8,60,79].

Blockchain-integrated healthcare architectures demonstrate how confidentiality and integrity can be combined but also highlight residual risks. In IoMT scenarios, blockchain-based systems are proposed to support secure sharing of medical records across institutions, improving disease prediction and remote monitoring while maintaining immutable logs of access and analysis [79]. However, transaction data remain accessible on public ledgers, and deanonymization attacks on Ethereum can link pseudonymous addresses to real-world identities; differential privacy has therefore been advocated to perturb stored or queried data by adding calibrated noise, with smart contracts enabling developers to embed noise injection strategies that balance privacy and utility for each application [60]. In remote patient monitoring, data breaches in medical companies have risen dramatically, and centralized IoT eHealth architectures that maintain patient data on single servers are vulnerable to ransomware, man-in-the-middle, and insider attacks, underscoring the need for distributed, privacy-preserving authentication and storage designs that restrict who can access behavioral and physiological data [8,79]. Ethical analyses of continuous authentication emphasize that confidentiality is not only a cryptographic problem but also a question of consent, transparency, and governance. Persistent sensor deployment for behavioral and contextual monitoring may infringe individual privacy rights when users are unaware of the extent of data collection and subsequent usage for authentication and profiling [4]. Research on continuous authentication explicitly warns that users may not understand how

far context-aware mechanisms capture and store GPS or other environmental information, which motivates transparent privacy policies and robust data protection measures that clearly define retention periods, access rights, and purposes for behavioral data processing [4,79]. Future directions therefore call for privacy-preserving authentication techniques, including zero-knowledge proof, Quantum Cryptography, Bloom filter-based template protection, cancelable biometrics, and fuzzy commitment strategies, together with ethical frameworks and interdisciplinary collaboration that establish principles for responsible collection, storage, and sharing of biometric and contextual information in continuous authentication systems [4,8,60,79].

6.2. Legislative Regulation

Regulatory frameworks around biometrics increasingly shape what continuous authentication systems may collect, how long data can be stored, and which rights individuals can exercise over their behavioral and physiological traces. Recent privacy laws such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) establish guidelines for using and storing biometric data, with GDPR explicitly classifying biometric data that allow or confirm unique identification as a special category of personal data under Article 9. This designation triggers heightened protection requirements and obliges controllers to incorporate measures such as template protection and careful selection of storage locations to conform with users' privacy expectations in biometric authentication scenarios [41].

Within GDPR, data subject rights codified in Chapter III have direct consequences for behavioral and continuous authentication deployments. Summaries of these provisions highlight that individuals must be informed about responsible parties using the data, the identity of the data controller, and the duration of storage, and must be able to request access, rectification, and erasure of their personal data. Additional rights cover data portability and protection against decisions based solely on automated processing, including profiling, which is particularly salient for systems that continuously infer user identity or soft biometric traits from behavior without explicit interaction. From a controller and processor standpoint, GDPR requires data protection by design, appropriate technical and organizational measures, security of processing, breach notification to supervisory authorities and subjects, and, where relevant, data protection impact assessments tailored to sensitive biometric processing.

Touch-based behavioral biometrics and keystroke dynamics illustrate how these regulatory expectations intersect with emerging profiling practices. Studies report an increasing focus on soft biometrics, such as age, gender, handedness, and one-hand versus two-hand typing, derived from typing and signature patterns. While such profiling can benefit domains like e-commerce through targeted advertising, it is described as intrinsic to behavioral schemes and therefore subject to GDPR's constraints on automated profiling and special category data handling. Commentaries stress that users should be aware of rights including erasure and portability with respect to their biometric profiles, and argue that data protection methods specific to behavioral biometric attributes remain underdeveloped and require further in depth research to meet regulatory standards [32].

Continuous authentication based on behavioral and contextual data must also align with broader privacy obligations around consent and transparency. Surveys of continuous authentication warn that users may be unaware of the extent to which biometric and contextual information is collected and reused for authentication, and call for transparent privacy policies and robust data protection measures. GDPR's emphasis on informed consent for obtaining and processing biometric data reinforces these concerns, since behavioral schemes often operate implicitly without active user input. Evidence therefore underscores

the need to adapt continuous monitoring architectures to respect legal requirements on notice, consent, and purpose limitation when capturing routine touch, motion, or usage traces over extended periods [4,32,41].

Blockchain-based smart home and IoT architectures introduce additional regulatory challenges regarding confidentiality and identifiability of transaction data. In public blockchains such as Ethereum, transactions are recorded in a distributed ledger accessible to all participants; although cryptographic mechanisms protect integrity, transaction details are broadcast and remain uniformly stored across the network. Analyses explain that public keys function as user identifiers, so linking techniques allow adversaries to deanonymize users by correlating on-chain data with external information, compromising personal privacy. Deanonymization attacks on Ethereum are specifically identified as a major risk, prompting calls for stronger privacy preservation strategies in blockchain-based applications that handle personal or behavioral data in home or IoT settings [60].

Researchers investigating privacy-preserving mechanisms for blockchain-based smart homes therefore propose strategies that align more closely with emerging data protection principles. Anonymization and two-level anonymity schemes are cited as initial approaches, but differential privacy is presented as a promising solution for ledger data. Differential privacy adds calibrated noise either to stored records or to query evaluations on the public ledger, with smart contracts allowing developers to embed noise injection tailored to privacy and utility requirements. This randomness aims to prevent adversaries from inferring individual behaviors or transaction histories from blockchain data, thereby addressing regulatory concerns around data minimization and identifiability in decentralized infrastructures that might otherwise expose detailed usage patterns of connected devices [60].

Smart grid and smart home environments further illustrate how regulatory-style requirements for access control and accountability translate into technical design choices. In B5G-enabled smart grids, authors argue that the heterogeneity of technologies and stakeholders, including prosumers, governments, utilities, and regulators, mandates hybrid access control mechanisms such as role-based and attribute-based models. Policy enforcement points in households and substations issue requests to policy decision points operated by aggregators or utilities, with all elements acting as policy information points that collect contextual information to enrich authorization decisions. Ubiquitous context awareness processes are proposed to permanently analyze infrastructure state so that authorization components can improve decision making and refine policies in real time, meeting security and timeliness requirements across regulatory and operational domains [19].

IoT-based smart home deployments confront similar governance demands but in more resource-constrained settings. Centralized access control architectures are criticized for communication overhead, single points of failure, and difficulties in regulating access to desired resources across heterogeneous devices. Distributed access control networks using blockchain are advocated as a way to provide trustworthy, dependable policies that can withstand malicious attacks, with nodes agreeing on rights assignment, policy rules, and verification outcomes. At the same time, authors caution that blockchain's distributed nature and public visibility raise unresolved issues regarding data preservation and transaction privacy, again motivating privacy enhancements consistent with data protection expectations, including improved anonymization and differential privacy [60]. Taken together, these analyses show that legislative and regulatory frameworks such as GDPR and CCPA already exert significant pressure on the design of continuous and behavioral biometric authentication, on blockchain-based infrastructures that may store related data, and on cyber-physical environments such as smart grids and smart homes.

6.3. Transparency and User Consent

Continuous authentication architectures frequently capture biometric, behavioral, and contextual signals in a largely invisible manner, which makes transparency and user consent central ethical requirements. Survey work on continuous authentication explicitly warns that continuous monitoring of user activities through sensors may infringe individual privacy rights, especially when individuals are not aware of how extensively their data are collected and reused for authentication and profiling. Context-aware modes, for instance, those relying on GPS or other environmental information, intensify these concerns because users may not realize that location, time of day, and activity patterns are being logged on a continuous basis and incorporated into identity decisions [4]. Reviews of sensor-based behavioral biometrics similarly report that users voice privacy concerns about the use of their behavioral data for continuous authentication and prefer solutions that are both less resource consuming and more respectful of personal data handling [4,9].

Regulatory analyses connect these transparency deficits directly to legal obligations. Under frameworks such as the General Data Protection Regulation, biometric data used for unique identification fall into a special category of personal data, and summaries highlight that subjects must be informed about who is using their data, the identity of the data controller, storage duration, and processing purposes. Individuals must also be able to request access, rectification, and erasure of stored biometric information, and enjoy protection against decisions based solely on automated processing and profiling. Work on touch-based behavioral biometrics notes that profiling of soft attributes such as age, gender, and handedness from typing and signature patterns has become more prominent, which strengthens the argument that users should be aware of these practices and of their associated rights to erasure and portability. Practical studies of keystroke and touch-based authentication indicate that systems routinely record very fine-grained timing and motion features, for example, dwell time, flight time, and accelerometer or gyroscope signals, and these low-level traces are commonly used to train advanced models including recurrent networks optimized with meta-heuristic techniques. The fact that researchers explicitly rely on such granular traces to boost classification accuracy may suggest that transparency about exactly which low-level signals are recorded, and for how long, is essential for meaningful user consent [80]. These commentaries further argue that data protection methods specific to behavioral biometric attributes are still immature and require deeper research to align with regulatory expectations on transparency and data subject control [32,41].

From a system design standpoint, several surveys underline that privacy policies and data protection mechanisms must be made explicit rather than implicit. Analyses of continuous authentication emphasize that users may be unaware of the extent to which biometric and contextual information is collected and utilized for security purposes, and call for clear, accessible privacy policies that describe collection, storage, retention, and sharing in concrete terms. The same work links transparency to consent, arguing that deployment of context-aware modes must include explicit mechanisms for obtaining user consent and for empowering individuals to control how their data are used. Proposed ethical frameworks for continuous authentication stress that developers and researchers should articulate principles governing collection, usage, and storage of biometric and contextual data, and that mechanisms for informed consent and user control are crucial to uphold autonomy and privacy rights in practice [4].

In blockchain-based smart home and IoT scenarios, transparency has a dual character. Public ledgers such as Ethereum inherently expose transaction data to all participants; cryptographic techniques protect integrity but do not in themselves guarantee confidentiality. Analyses explain that public keys function as user identifiers and that deanonymization attacks can link on-chain histories with external databases, thereby revealing the identities

of users behind specific transactions. This visibility implies that users should be clearly informed that interaction with blockchain-based access control or authentication mechanisms can leave permanent, globally accessible traces, and that simple use of hashes or public key cryptography does not equate to anonymity. To mitigate these issues, researchers discuss differential privacy-based perturbation that adds calibrated noise to stored records or query evaluations, and note that smart contracts can embed such noise injection strategies. Even in these enhanced designs, the need for transparency about how much noise is added, what privacy guarantees are provided, and how accuracy is affected remains critical for meaningful consent [60].

Federated and edge learning frameworks in healthcare-oriented IoT further demonstrate how architectural choices interact with user expectations. In edge computing-enabled smart healthcare, federated learning keeps patient data on local devices while sending only model updates to aggregation servers, with the stated aim of reducing backbone load and improving privacy. Nevertheless, discussions in this domain acknowledge unresolved questions about whether such infrastructures sufficiently protect the security and privacy of clinical data, and identify privacy, security, level of trust, traceability, and accountability as open challenges. These acknowledgements imply that participants in federated healthcare analytics should be informed not only that their raw data remain local, but also that model parameters transmitted for aggregation may still carry residual information about their medical records, and that concrete guarantees about data reconstruction resistance have yet to be fully established [69].

Ethical recommendations for continuous authentication consistently return to transparency and consent as areas demanding structured guidance. One survey proposes that future work should develop ethical frameworks and guidelines that explicitly govern collection, usage, and storage of biometric and contextual data, and additionally calls for mechanisms that allow users to control their data. Suggested directions include empowering individuals to decide which modalities are active, when data may be shared with cloud services, and under what conditions templates can be revoked or replaced. Interdisciplinary collaboration between computer science, psychology, ethics, and law is presented as essential for designing consent interactions that are understandable, non-coercive, and sensitive to user perceptions of risk, while still supporting robust authentication [4].

Privacy-preserving cryptographic techniques are positioned as technical enablers for more trustworthy consent relationships. Bloom filters and cancelable biometrics can deliver irreversible and unlinkable templates, making it easier to state and honor commitments that biometric samples will not be reconstructable if databases are compromised. Zero-knowledge-proof-based protocols allow authentication without revealing cryptographic keys, metadata, or explicit user authentication information to servers, which supports architectures where service providers can honestly claim minimal visibility into users' biometric secrets. Ethical analyses argue that such technical measures must be accompanied by transparent communication so that users understand what is protected, what remains exposed, and what rights they have if they later withdraw consent or if systems are breached [4,9,60,69].

7. Conclusions and Future Work

This review synthesizes findings on continuous mobile authentication that leverage behavioral biometrics, multimodal fusion, and integration with intelligent personal agents, wearables, and IoT/edge infrastructures. Behavioral signals such as keystroke timing, touch and swipe dynamics, gait, and motion patterns provide effective means for passive, ongoing verification that complements physiological traits used at initial entry. When combined with contextual cues, multimodal systems achieve improved discrimination and

greater operational flexibility, while context-aware adaptation can reduce intrusive user interactions by invoking stricter checks only when anomalous patterns are detected.

Tradeoffs between security and usability remain central. Operating thresholds must balance False Acceptance and False Rejection rates in line with application risk and user tolerance. Scalar summaries such as Equal Error Rate and HTER are useful for comparison but do not capture temporal drift, cross-device variability, energy impact, or the user experience costs of repeated re-authentication. Practical deployments therefore require dynamic thresholding, per-user adaptation, session-level aggregation, and clearly defined fallback behaviors that limit disruption while preserving protection.

Adversarial and systemic threats affect both sensing modalities and learning pipelines. Generative and imitation attacks can substantially degrade performance in touch and keystroke systems; gait and cardiac modalities are also vulnerable when representative traces are available to adversaries. Model-level threats, including membership inference, reconstruction, extraction, and poisoning, are realistic in cloud and federated settings. Protecting template confidentiality and limiting server visibility are therefore essential. Cancelable templates, Bloom filter encodings, zero-knowledge proof protocols, secure aggregation, and local or federated processing present concrete mitigation strategies. Architectures that partition computation between device, edge, and cloud can reduce data exposure but must include controls against malicious or curious insiders.

Resource constraints on mobile and wearable platforms impose further design requirements. Continuous sensing and feature computation increase energy consumption and may reduce device utility if not managed. Lightweight models, selective sampling, on-device aggregation, energy-aware modality selection informed by current activity, and edge-assisted inference can reduce communication and runtime overhead. Trusted Execution Environments and personal agents that enforce local policies and provenance controls help limit unnecessary data sharing without overloading remote services.

Advances in machine learning expand capabilities but introduce new challenges. Classical classifiers remain attractive where labeled data are limited and latency constraints are tight; deep architectures, RNNs, CNNs, and transformers, deliver lower error rates on large, curated datasets and support end-to-end representation learning. Yet deep models may generalize poorly across domains, be sensitive to adversarial inputs, and raise explainability concerns in security-critical flows. Emerging approaches such as self-supervised contrastive encoders, privacy-aware federated strategies, and hybrid handcrafted plus learned feature designs offer promising paths to increase resilience while reducing data exposure.

Evaluation practice and dataset availability are notable weaknesses. Many studies report strong within-session or same-device performance but omit assessments of long-term stability, cross-device transfer, or resilience to unseen attack types. Public datasets vary widely in size, demographic coverage, and acquisition realism, and there is limited agreement on standardized benchmarks or reporting protocols. Adopting transparent review protocols, unified evaluation procedures that include adversarial testing and longitudinal enrollment scenarios, and explicit measurements of energy and latency will improve comparability and accelerate maturation.

Legal, ethical, and user-centered concerns must guide deployment. Continuous collection of behavioral and contextual traces implicates privacy, consent, and rights such as access, rectification and erasure under frameworks like GDPR and CCPA. Transparent notices, granular consent mechanisms, revocable and unlinkable template schemes, and explainable decision pathways increase acceptability. In ledger-based or federated architectures, attention to transaction visibility, deanonymization risk, and model update privacy is required; differential privacy and selective on-chain/off-chain strategies can help reconcile auditability with confidentiality.

Recommended directions for research and engineering include: developing adaptive multimodal systems that select modalities by activity and device state; embedding privacy-preserving cryptographic and template protection mechanisms into end-to-end pipelines; designing energy-conscious sensing and inference strategies suitable for wearables and edge nodes; strengthening defenses and standardized evaluations against adversarial and insider threats; and expanding public, diverse, and longitudinal datasets with realistic attack scenarios. Interdisciplinary collaboration that combines algorithmic advances with human factors, regulatory compliance, and ethical guidance will be essential to produce continuous authentication solutions that are effective, acceptable to users, and practicable in mobile and assistant-enriched environments.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/technologies14070451/s1>, Search_Strategy.txt; Quality_Assessment_Scores.xlsx [81].

Author Contributions: Conceptualization, M.G. and V.S.; methodology, M.G.; formal analysis, M.G.; investigation, M.G., A.K., Y.C. and A.A.; resources, V.S.; data curation, M.G.; writing—original draft preparation, M.G.; writing—review and editing, M.G., A.K., Y.C., A.A. and V.S.; visualization, M.G.; supervision, V.S.; project administration, V.S.; funding acquisition, V.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan, grant number BR24993072.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset generated and analyzed during the current study, including the full list of included articles, bibliometric records, screening logs, and data extraction tables, is publicly available in the GitHub repository: <https://github.com/madigali/Behavioral-Biometric-Continuous-Authentication-for-Mobile-Devices-with-an-Intelligent-Personal-Agent/tree/main> (accessed on 14 July 2026). All data were derived from publicly accessible peer-reviewed publications.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (OpenAI, GPT-5.2 version) exclusively for language translation and linguistic refinement. All scientific ideas, methodology, analysis, system design, and conclusions were developed independently by the authors. The authors have carefully reviewed and edited all generated text and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Skalkos, A.; Stylios, I.; Karyda, M.; Kokolakis, S. Users' Privacy Attitudes towards the Use of Behavioral Biometrics Continuous Authentication (BBCA) Technologies: A Protection Motivation Theory Approach. *J. Cybersecur. Priv.* **2021**, *1*, 743–766. [CrossRef]
2. Hasan, S.S.U.; Ghani, A.; Daud, A.; Akbar, H.; Khan, M.F. A Review on Secure Authentication Mechanisms for Mobile Security. *Sensors* **2025**, *25*, 700. [CrossRef] [PubMed]
3. Abuhamad, M.; Abusnaina, A.; Nyang, D.; Mohaisen, D. Sensor-Based Continuous Authentication of Smartphones' Users Using Behavioral Biometrics: A Contemporary Survey. *IEEE Internet Things J.* **2021**, *8*, 65–84. [CrossRef]
4. Ayeswarya, S.; Singh, K.J. A Comprehensive Review on Secure Biometric-Based Continuous Authentication and User Profiling. *IEEE Access* **2024**, *12*, 82996–83021. [CrossRef]
5. de Barcelos Silva, A.; Gomes, M.M.; da Costa, C.A.; da Rosa Righi, R.; Barbosa, J.L.V.; Pessin, G.; Doncker, G.D.; Federizzi, G. Intelligent Personal Assistants: A Systematic Literature Review. *Expert Syst. Appl.* **2020**, *147*, 113193. [CrossRef]
6. Hakeem, S.A.A.; Hussein, H.H.; Kim, H. Security Requirements and Challenges of 6G Technologies and Applications. *Sensors* **2022**, *22*, 1969. [CrossRef] [PubMed]

7. Alqahtani, F.; Al-Makhadmeh, Z.; Tolba, A.; Said, O. TBM: A trust-based monitoring security scheme to improve the service authentication in the Internet of Things communications. *Comput. Commun.* **2019**, *150*, 216–225. [[CrossRef](#)]
8. Deebak, B.D.; Hwang, S.O. Federated Learning-Based Lightweight Two-Factor Authentication Framework with Privacy Preservation for Mobile Sink in the Social IoMT. *Electronics* **2023**, *12*, 1250. [[CrossRef](#)]
9. Baig, A.F.; Eskeland, S. Security, Privacy, and Usability in Continuous Authentication: A Survey. *Sensors* **2021**, *21*, 5967. [[CrossRef](#)] [[PubMed](#)]
10. Finnegan, O.L.; White, J.W., III; Armstrong, B.; Adams, E.L.; Burkart, S.; Beets, M.W.; Nelakuditi, S.; Willis, E.A.; von Klinggraff, L.; Parker, H.; et al. The utility of behavioral biometrics in user authentication and demographic characteristic detection: A scoping review. *Syst. Rev.* **2024**, *13*, 61. [[CrossRef](#)] [[PubMed](#)]
11. Roy, S.; Pradhan, J.; Kumar, A.; Adhikary, D.R.D.; Roy, U.; Sinha, D.; Pal, R.K. A Systematic Literature Review on Latest Keystroke Dynamics Based Models. *IEEE Access* **2022**, *10*, 92192–92236. [[CrossRef](#)]
12. Khan, Z.F.; Alotaibi, S.R. Design and Implementation of a Computerized User Authentication System for E-Learning. *Int. J. Emerg. Technol.* **2020**, *15*, 4–18. [[CrossRef](#)]
13. Wang, C.; Wang, Y.; Chen, Y.; Liu, H.; Liu, J. User Authentication on Mobile Devices: Approaches, Threats and Trends. *Comput. Netw.* **2020**, *170*, 107118. [[CrossRef](#)]
14. Verma, A.; Moghaddam, V.; Anwar, A. Data-Driven Behavioural Biometrics for Continuous and Adaptive User Verification Using Smartphone and Smartwatch. *Sustainability* **2022**, *14*, 7362. [[CrossRef](#)]
15. Li, G.; Sato, H. Sensing In-Air Signature Motions Using Smartwatch: A High-Precision Approach of Behavioral Authentication. *IEEE Access* **2022**, *10*, 57865–57879. [[CrossRef](#)]
16. Khairnar, S.; Gite, S.; Kotecha, K.; Thepade, S.D. Face Liveness Detection Using Artificial Intelligence Techniques: A Systematic Literature Review and Future Directions. *Big Data Cogn. Comput.* **2023**, *7*, 37. [[CrossRef](#)]
17. Kollu, V.N.; Janarthanan, V.; Karupusamy, M.; Ramachandran, M. Cloud-Based Smart Contract Analysis in FinTech Using IoT-Integrated Federated Learning in Intrusion Detection. *Data* **2023**, *8*, 83. [[CrossRef](#)]
18. Kim, M.; Oh, I.; Yim, K.; Sahlabadi, M.; Shukur, Z. Security of 6G-Enabled Vehicle-to-Everything Communication in Emerging Federated Learning and Blockchain Technologies. *IEEE Access* **2024**, *12*, 33972–34001. [[CrossRef](#)]
19. Lopez, J.; Rubio, J.E.; Alcaraz, C. Digital Twins for Intelligent Authorization in the B5G-enabled Smart Grid. *IEEE Wirel. Commun.* **2022**, *28*, 48–55. [[CrossRef](#)]
20. Ağca, M.A.; Faye, S.; Khadraoui, D. Trusted Distributed Artificial Intelligence (TDAI). *IEEE Access* **2023**, *11*, 113307–113323. [[CrossRef](#)]
21. Hirschprung, R.S.; Alkoby, S. A Game Theory Approach for Assisting Humans in Online Information-Sharing. *Information* **2022**, *13*, 183. [[CrossRef](#)]
22. Mekruksavanich, S.; Jitpattanakul, A. Deep Learning Approaches for Continuous Authentication Based on Activity Patterns Using Mobile Sensing. *Sensors* **2021**, *21*, 7519. [[CrossRef](#)] [[PubMed](#)]
23. Stragapede, G.; Delgado-Santos, P.; Tolosana, R.; Vera-Rodriguez, R.; Guest, R.; Morales, A. Mobile Keystroke Biometrics Using Transformers. In *Proceedings of the 2023 IEEE 17th International Conference on Automatic Face and Gesture Recognition (FG)*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6. [[CrossRef](#)]
24. Rocha, R.; Carneiro, D.; Novais, P. Continuous authentication with a focus on explainability. *Neurocomputing* **2020**, *423*, 697–702.
25. Michailidis, E.T.; Vouyioukas, D. A Review on Software-Based and Hardware-Based Authentication Mechanisms for the Internet of Drones. *Drones* **2022**, *6*, 41. [[CrossRef](#)]
26. Fan, M.; Zhang, Z.; Li, Z.; Sun, G.; Yu, H.; Guizani, M. Blockchain-based Decentralized and Lightweight Anonymous Authentication for Federated Learning. *IEEE Trans. Veh. Technol.* **2023**, *72*, 12075–12086. [[CrossRef](#)]
27. Şahin, D.Ö.; Akleyek, S.; Kiliç, E. LinRegDroid: Detection of Android Malware Using Multiple Linear Regression Models-Based Classifiers. *IEEE Access* **2022**, *10*, 14246–14259. [[CrossRef](#)]
28. Wang, C.; Li, Z.; Sarpong, B. Multimodal adaptive identity-recognition algorithm fused with gait perception. *Big Data Min. Anal.* **2021**, *4*, 223–232. [[CrossRef](#)]
29. Hernández Sánchez, S.; Fernández Pozo, R.; Hernández Gómez, L.A. Driver Identification and Verification from Smartphone Accelerometers Using Deep Neural Networks. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 97–109. [[CrossRef](#)]
30. de Marcos, L.; Martínez-Herráiz, J.J.; Junquera-Sánchez, J.; Cilleruelo, C.; Pages-Arévalo, C. Comparing Machine Learning Classifiers for Continuous Authentication on Mobile Devices by Keystroke Dynamics. *Electronics* **2021**, *10*, 1622. [[CrossRef](#)]
31. Stylios, I.; Kokolakis, S.; Thanou, O.; Chatzis, S. Behavioral biometrics & continuous user authentication on mobile devices: A survey. *Inf. Fusion* **2021**, *66*, 76–99. [[CrossRef](#)]
32. Ellavarason, E.; Guest, R.; Deravi, F.; Sanchez-Reillo, R.; Corsetti, B. Touch-Dynamics Based Behavioural Biometrics on Mobile Devices: A Review from a Usability and Performance Perspective. *ACM Comput. Surv.* **2020**, *53*, 120:1–120:36. [[CrossRef](#)]
33. Bakkar, M.; Alazab, A. Designing Security Intelligent Agent for Petrol Theft Prevention. In *Proceedings of the 2019 Cybersecurity and Cyberforensics Conference (CCC)*; IEEE: Piscataway, NJ, USA, 2019; pp. 123–128. [[CrossRef](#)]

34. Deng, Z.; Guo, Y.; Han, C.; Ma, W.; Xiong, J.; Wen, S.; Xiang, Y. AI Agents Under Threat: A Survey of Key Security Challenges and Future Pathways. *ACM Comput. Surv.* **2025**, *57*, 182. [\[CrossRef\]](#)
35. Pryor, L.; Dave, R.; Seliya, J.; Sowell-Boone, E. Machine Learning Algorithms in User Authentication Schemes. In *Proceedings of the 2021 IEEE International Conference on Electrical, Computer and Energy Technologies (ICECET)*; IEEE: Piscataway, NJ, USA, 2021; pp. 1–6. [\[CrossRef\]](#)
36. Kamble, M.R.; Sailor, H.B.; Patil, H.A.; Li, H. Advances in anti-spoofing: From the perspective of ASVspoof challenges. *Apsipa Trans. Signal Inf. Process.* **2020**, *9*, e2. [\[CrossRef\]](#)
37. Kokal, S.; Vanamala, M.; Dave, R. Deep Learning and Machine Learning, Better Together Than Apart: A Review on Biometrics Mobile Authentication. *J. Cybersecur. Priv.* **2023**, *3*, 227–258. [\[CrossRef\]](#)
38. Ray-Dowling, A.; Hou, D.; Schuckers, S.C. Stationary mobile behavioral biometrics: A survey. *Comput. Secur.* **2023**, *128*, 103184. [\[CrossRef\]](#)
39. Gupta, C.; Johri, I.; Srinivasan, K.; Hu, Y.C.; Qaisar, S.M.; Huang, K.Y. A Systematic Review on Machine Learning and Deep Learning Models for Electronic Information Security in Mobile Networks. *Sensors* **2022**, *22*, 2017. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Acien, A.; Morales, A.; Monaco, J.V.; Vera-Rodriguez, R.; Fierrez, J. TypeNet: Deep Learning Keystroke Biometrics. *IEEE Trans. Biom. Behav. Identity Sci.* **2022**, *4*, 57–70. [\[CrossRef\]](#)
41. Buriro, A.; Gupta, S.; Yautsiukhin, A.; Crispo, B. Risk-Driven Behavioral Biometric-based One-Shot-cum-Continuous User Authentication Scheme. *J. Signal Process. Syst.* **2021**, *93*, 989–1006. [\[CrossRef\]](#)
42. Klein, I. Smartphone Location Recognition: A Deep Learning-Based Approach. *Sensors* **2020**, *20*, 214. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Kececi, A.; Yildirak, A.; Ozyazici, K.; Ayluctarhan, G.; Agbulut, O.; Zincir, I. Gait Recognition Using Siamese Neural Networks with Wearable Sensors. *Eng. Sci. Technol. Int. J.* **2020**, *23*, 931–937. [\[CrossRef\]](#)
44. Huang, H.; Zhou, P.; Li, Y.; Sun, F. A Lightweight Attention-Based CNN Model for Efficient Gait Recognition with Wearable IMU Sensors. *Sensors* **2021**, *21*, 2866. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Mekruksavanich, S.; Jitpattanakul, A. Biometric User Identification Based on Human Activity Recognition Using Wearable Sensors: An Experiment Using Deep Learning Models. *Electronics* **2021**, *10*, 308. [\[CrossRef\]](#)
46. Wang, Y.; Zhang, X.; Hu, H. Continuous User Authentication on Multiple Smart Devices. *Information* **2023**, *14*, 274. [\[CrossRef\]](#)
47. Zou, Q.; Wang, Y.; Wang, Q.; Zhao, Y.; Li, Q. Deep Learning-Based Gait Recognition Using Smartphones in the Wild. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 3197–3212. [\[CrossRef\]](#)
48. De Marsico, M.; Palermo, A. User gait biometrics in smart ambient applications through wearable accelerometer signals. *J. Ambient. Intell. Humaniz. Comput.* **2024**, *15*, 2967–2979. [\[CrossRef\]](#)
49. Estrela, P.M.A.B.; Albuquerque, R.d.O.; Amaral, D.M.; Giozza, W.F.; Júnior, R.T.d.S. A Framework for Continuous Authentication Based on Touch Dynamics Biometrics for Mobile Banking Applications. *Sensors* **2021**, *21*, 4212. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Huang, A.; Gao, S.; Chen, J.; Xu, L.; Nathan, A. High Security User Authentication Enabled by Piezoelectric Keystroke Dynamics and Machine Learning. *IEEE Sens. J.* **2020**, *20*, 13037–13046. [\[CrossRef\]](#)
51. Li, W.; Tan, J.; Meng, W.; Wang, Y. A swipe-based unlocking mechanism with supervised learning on smartphones: Design and evaluation. *J. Netw. Comput. Appl.* **2020**, *165*, 102687. [\[CrossRef\]](#)
52. Alzahrani, S.; Alderaan, J.; Alatawi, D.; Alotaibi, B. Continuous Mobile User Authentication Using a Hybrid CNN-Bi-LSTM Approach. *Comput. Mater. Contin.* **2023**, *75*, 651–667. [\[CrossRef\]](#)
53. Incel, Ö.D.; Günay, S.; Akan, Y.; Barlas, Y.; Basar, O.E.; Alptekin, G.I.; Isbilen, M. DAKOTA: Sensor and Touch Screen-Based Continuous Authentication on a Mobile Banking Application. *IEEE Access* **2021**, *9*, 38943–38960. [\[CrossRef\]](#)
54. Ackerson, J.M.; Dave, R.; Seliya, N. Applications of Recurrent Neural Network for Biometric Authentication & Anomaly Detection. *Information* **2021**, *12*, 272. [\[CrossRef\]](#)
55. Kim, J.; Kang, P. Freely typed keystroke dynamics-based user authentication for mobile devices based on heterogeneous features. *Pattern Recognit.* **2020**, *108*, 107556. [\[CrossRef\]](#)
56. Kumar, S.; Rani, S.; Jain, A.; Verma, C.; Raboaca, M.S.; Illés, Z.; Neagu, B.C. Face Spoofing, Age, Gender and Facial Expression Recognition Using Advance Neural Network Architecture-Based Biometric System. *Sensors* **2022**, *22*, 5160. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Tapia, J.E.; Gonzalez, S.; Busch, C. Iris Liveness Detection Using a Cascade of Dedicated Deep Learning Networks. *IEEE Trans. Inf. Forensics Secur.* **2022**, *17*, 42–52. [\[CrossRef\]](#)
58. Yu, C.; Cheng, Z.; Cui, H.; Gao, Y.; Luo, Z.; Wang, Y.; Zheng, H.; Zhao, Y. A Survey on Agent Workflow—Status and Future. In *Proceedings of the 2025 8th International Conference on Artificial Intelligence and Big Data (ICAIBD)*; IEEE: Piscataway, NJ, USA, 2025; pp. 770–781. [\[CrossRef\]](#)
59. Ryu, R.; Yeom, S.; Herbert, D.; Dermoudy, J. The design and evaluation of adaptive biometric authentication systems: Current status, challenges and future direction. *ICT Express* **2023**, *9*, 1183–1197. [\[CrossRef\]](#)
60. Qashlan, A.; Nanda, P.; He, X.; Mohanty, M. Privacy-Preserving Mechanism in Smart Home Using Blockchain. *IEEE Access* **2021**, *9*, 103651–103669. [\[CrossRef\]](#)

61. Zhao, J.; Hu, H.; Huang, F.; Guo, Y.; Liao, L. Cross-Domain Anonymous Authentication for Vehicular Networks Using Blockchain. *Electronics* **2023**, *12*, 1812. [\[CrossRef\]](#)
62. Kokal, S.; Pryor, L.; Dave, R. Exploration of Machine Learning Classification Models Used for Behavioral Biometrics Authentication. In *Proceedings of the 2022 8th International Conference on Computer Technology Applications*; Association for Computing Machinery: New York, NY, USA, 2022; pp. 176–182. [\[CrossRef\]](#)
63. Shen, Z.; Li, S.; Zhao, X.; Zou, J. MMAuth: A Continuous Authentication Framework on Smartphones Using Multiple Modalities. *IEEE Trans. Inf. Forensics Secur.* **2022**, *17*, 1450–1465. [\[CrossRef\]](#)
64. Gomez-Alanis, A.; Gonzalez-Lopez, J.A.; Peinado, A.M. A Kernel Density Estimation Based Loss Function and its Application to ASV-Spoofing Detection. *IEEE Access* **2020**, *8*, 108530–108543. [\[CrossRef\]](#)
65. Uliyan, D.M.; Sadeghi, S.; Jalab, H.A. Anti-spoofing method for fingerprint recognition using patch based deep learning machine. *Eng. Sci. Technol. Int. J.* **2019**, *23*, 264–273. [\[CrossRef\]](#)
66. Ntizikira, E.; Lei, W.; Alblehai, F.; Saleem, K.; Lodhi, M.A. Secure and Privacy-Preserving Intrusion Detection and Prevention in the Internet of Unmanned Aerial Vehicles. *Sensors* **2023**, *23*, 8077. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Rizk, H.; Elmogy, A. Self-Supervised WiFi-Based Identity Recognition in Multi-User Smart Environments. *Sensors* **2025**, *25*, 3108. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Bidgoly, A.J.; Bidgoly, H.J.; Arezoumand, Z. A Survey on Methods and Challenges in EEG Based Authentication. *Comput. Secur.* **2020**, *93*, 101788. [\[CrossRef\]](#)
69. Abreha, H.G.; Hayajneh, M.; Serhani, M.A. Federated Learning in Edge Computing: A Systematic Survey. *Sensors* **2022**, *22*, 450. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Li, Z.; Shi, C.; Xie, Y.; Liu, J.; Yuan, B.; Chen, Y. Practical Adversarial Attacks Against Speaker Recognition Systems. In *Proceedings of the 21st International Workshop on Mobile Computing Systems and Applications*; Association for Computing Machinery: New York, NY, USA, 2020; p. 6. [\[CrossRef\]](#)
71. Sun, F.; Zang, W.; Gravina, R.; Fortino, G.; Li, Y. Gait-based identification for elderly users in wearable healthcare systems. *Inf. Fusion* **2020**, *53*, 134–144. [\[CrossRef\]](#)
72. Mkpa, A.; Chin, J.; Winckles, A. Holistic Blockchain Approach to Foster Trust, Privacy and Security in IoT Based Ambient Assisted Living Environment. In *Proceedings of the 2019 15th International Conference on Intelligent Environments (IE)*; IEEE: Piscataway, NJ, USA, 2019; pp. 52–55. [\[CrossRef\]](#)
73. Din, I.U.; Khan, K.H.; Almogren, A.; Zareei, M.; Díaz, J.A.P. Securing the Metaverse: A Blockchain-Enabled Zero-Trust Architecture for Virtual Environments. *IEEE Access* **2024**, *12*, 92337–92347. [\[CrossRef\]](#)
74. Awadallah, A.; Eledlebi, K.; Zemerly, M.J.; Puthal, D.; Damiani, E.; Taha, K.; Kim, T.-Y.; Yoo, P.D.; Choo, K.-K.R.; Yim, M.-S.; et al. Artificial Intelligence-Based Cybersecurity for the Metaverse: Research Challenges and Opportunities. *IEEE Commun. Surv. Tutor.* **2025**, *27*, 1008–1052. [\[CrossRef\]](#)
75. Sufi, F. A New AI-Based Semantic Cyber Intelligence Agent. *Future Internet* **2023**, *15*, 231. [\[CrossRef\]](#)
76. Choi, M.; Lee, S.; Jo, M.; Shin, J.S. Keystroke Dynamics-Based Authentication Using Unique Keypad. *Sensors* **2021**, *21*, 2242. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Ming, Z.; Visani, M.; Luqman, M.M.; Burie, J.C. A Survey on Anti-Spoofing Methods for Facial Recognition with RGB Cameras of Generic Consumer Devices. *J. Imaging* **2020**, *6*, 139. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Kim, D.I.; Lee, S.; Shin, J.S. A New Feature Scoring Method in Keystroke Dynamics-Based User Authentications. *IEEE Access* **2020**, *8*, 27901–27914. [\[CrossRef\]](#)
79. Rahman, A.; Wadud, M.A.H.; Islam, M.J.; Kundu, D.; Bhuiyan, T.M.A.U.H.; Muhammad, G.; Ali, Z. Internet of medical things and blockchain-enabled patient-centric agent through SDN for remote patient monitoring in 5G network. *Sci. Rep.* **2024**, *14*, 5297. [\[CrossRef\]](#) [\[PubMed\]](#)
80. El-Kenawy, E.S.M.; Mirjalili, S.; Abdelhamid, A.A.; Ibrahim, A.; Khodadadi, N.; Eid, M.M. Meta-Heuristic Optimization and Keystroke Dynamics for Authentication of Smartphone Users. *Mathematics* **2022**, *10*, 2912. [\[CrossRef\]](#)
81. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.