computers

MDPI

Article

Context-Aware Decision Fusion for Multimodal Access Control Under Contradictory Biometric Evidence

Yasser Hmimou [1,2,*], Azedine Khiat [2], Hassna Bensag [2], Zineb Hidila [1] and Mohamed Tabaa [1,*]

[1] Multidisciplinary Laboratory of Research and Innovation (LPRI), Moroccan School of Engineering Sciences (EMSI), Casablanca 20250, Morocco; z.hidila@emsi.ma
[2] Computing, Artificial Intelligence and Cyber Security Laboratory (2IACS), ENSET Mohammedia, Hassan II University of Casablanca, Casablanca 20360, Morocco; khiat@enset-media.ac.ma (A.K.); h.bensag@gmail.com (H.B.)
* Correspondence: y.hmimou@emsi.ma (Y.H.); m.tabaa@emsi.ma (M.T.)

Abstract

Access control systems rely increasingly on multimodal biometric and behavioral signals to enhance security and robustness against sophisticated attacks. However, when heterogeneous modalities provide conflicting evidence, such as valid biometric credentials accompanied by abnormal behavioral or acoustic patterns, traditional fusion strategies based on static thresholds or majority voting often fail, leading to false alarms or insecure authorization decisions. This paper addresses this critical limitation by proposing a contextual decision-making fusion framework designed to resolve conflicting multimodal evidence at the decision-making level. The proposed approach models access control as a decision-making problem in a context of uncertainty, where independent agents generate modality-specific evidence from authentication channels based on face, voice, and fingerprints. A centralized fusion mechanism integrates heterogeneous results using adaptive reliability weighting and contextual reasoning to resolve conflicts before operational decisions are made. Rather than treating each modality independently, the framework explicitly considers inconsistencies, uncertainties, and situational context when aggregating evidence. The framework is evaluated using public benchmarks, including VGGFace2, VoxCeleb2, and FVC2004, combined with controlled multimodal scenarios that induce conflicting evidence. Experimental results obtained under controlled contradiction scenarios show that the proposed fusion strategy reduces false alarms and improves decision consistency by approximately 18%. These results are interpreted within the scope of controlled multimodal simulations.

Keywords: multimodal biometric authentication; decision fusion; contradiction handling; context-aware systems; access control; internet of things; reliability weighting

Academic Editor: Paolo Bellavista

Received: 25 February 2026
Revised: 23 March 2026
Accepted: 24 March 2026
Published: 27 March 2026

1. Introduction

Biometric access control systems are central to modern security infrastructures, enabling reliable identity verification across a wide range of applications, including physical access control, digital authentication, and Internet of Things (IoT) environments. Early biometric systems relied primarily on unimodal features such as facial images or fingerprints, which, despite their convenience, remain vulnerable to sensor noise, environmental variability, and impersonation attacks. To overcome these limitations, multimodal biometric authentication has been extensively studied and widely adopted. It combines complementary biometric features to improve robustness and reliability [1–3]. Recent surveys and

large-scale studies confirm that multimodal approaches consistently outperform unimodal systems in real-world, unconstrained conditions [4,5], and are particularly well suited to distributed, IoT-based security scenarios [6].

Rapid advances in deep learning have further accelerated the development of multimodal biometric systems. Advanced neural architectures now allow for the extraction of discriminative features, attention-based fusion, and adaptive weighting between heterogeneous biometric modalities [7–9]. These techniques have been successfully applied to a wide range of biometric signals, including facial images, fingerprints, speech, gait, and behavioral characteristics [10–13]. As a result, multimodal authentication has become a key element for secure access control in smart environments, industrial IoT, and cyber-physical systems.

Despite these advances, the increasing complexity of multimodal authentication pipelines poses a major challenge that remains insufficiently addressed: the presence of conflicting biometric evidence at the time of decision-making. Under realistic operating conditions, different modalities can produce inconsistent or contradictory results. For example, a user may present valid facial and fingerprint credentials while exhibiting abnormal voice or behavioral patterns due to stress, illness, ambient noise, or sensor degradation. Existing access control frameworks and fusion strategies are generally designed with the implicit assumption that biometric modalities provide consistent evidence. When this hypothesis is challenged, conventional fusion mechanisms, such as static thresholding, weighted averaging, or majority voting, can yield high false-rejection rates or insecure authorization decisions [14–16].

Numerous previous studies have focused on improving multimodal fusion at the feature or score level, often by optimizing classification accuracy through designed aggregation rules or learning-based fusion schemes [17–19]. While these approaches demonstrate robust performance under nominal conditions, they rarely model uncertainty at the decision level or explicitly account for contradictions between modalities. In practice, dominant modalities may overshadow weaker but contextually significant signals, or noisy inputs may disproportionately influence the final decision. Such behavior is particularly problematic in security-critical access control scenarios, where false acceptances and false rejections have significant operational and safety consequences.

At the same time, adaptive and intelligent authentication mechanisms have been proposed to improve robustness and context awareness, particularly in IoT and distributed environments [20–23]. These solutions incorporate behavioral biometrics, adaptive authentication policies, explainable AI components, and AI-based security mechanisms [24,25]. Although these approaches improve flexibility and resilience, they focus primarily on improving the accuracy, scalability, or interpretability of authentication, rather than explicitly addressing the problem of decision fusion in the presence of conflicting multimodal evidence. Therefore, the question of how to reliably arbitrate access decisions when biometric signals disagree remains largely open.

This paper addresses this gap by reformulating multimodal access control as a decision fusion problem in a context of uncertainty, in which contradictions are treated as first-order phenomena rather than anomalies to be ignored. We propose a contextual decision fusion framework in which independent agents generate modality-specific evidence from facial, voice, and fingerprint-based authentication channels, while a centralized fusion mechanism resolves these heterogeneous results at the decision level. Unlike conventional fusion schemes, the proposed framework explicitly considers inconsistencies, uncertainties, and contextual reliability when aggregating evidence, enabling robust authorization decisions even in the presence of conflicting biometric signals.

It is important to note that this work focuses on the controlled analysis of contradiction-aware decision mechanisms rather than end-to-end multimodal system validation. This requires explicit control over modality disagreement patterns, which is difficult to achieve in naturally acquired synchronized datasets without additional annotation effort.

The proposed framework is evaluated using widely adopted public benchmarks, including VGGFace2, VoxCeleb2, and FVC2004, combined with controlled experimental scenarios specifically designed to induce multimodal conflicts [26–28]. Additional datasets are used to evaluate robustness under non-constrained conditions and noisy environments [29–31]. Experimental results demonstrate that conflict resolution at the decision level significantly reduces false alarms and improves decision consistency compared to conventional fusion baselines, while maintaining competitive authentication performance.

Unlike conventional multimodal fusion approaches, which implicitly assume consistency between biometric modalities and rely on static aggregation rules such as majority voting or score averaging, the proposed framework explicitly models contradictory biometric evidence at the decision level.

The novelty of this work lies in introducing a context-aware contradiction handling mechanism that integrates reliability gating, contextual signal evaluation, and utility-based decision arbitration. This design enables the system to reason about disagreements between modalities rather than treating them as noise, thereby improving robustness in realistic operating conditions where biometric signals may become unreliable or inconsistent.

The main contributions of this article are threefold. First, we identify and formalize the problem of conflicting multimodal biometric evidence in access control systems, highlighting the limitations of existing fusion strategies. Second, we propose a context-aware decision-making fusion framework that explicitly resolves inconsistencies at the decision level rather than relying on static aggregation rules. Finally, we provide a comprehensive experimental evaluation on public datasets, demonstrating the effectiveness of the proposed approach in improving reliability and stability in security-critical access control environments. Unlike existing fusion approaches that implicitly redistribute conflicting evidence, the proposed framework explicitly models contradiction as a first-class decision variable, transforming multimodal fusion into a risk-aware decision-making process rather than a simple aggregation problem.

The remainder of this article is structured as follows. Section 2 reviews related work on multimodal biometric authentication and decision fusion, with a particular focus on the limitations of existing fusion strategies when dealing with conflicting evidence. Section 3 formalizes the problem of decision-level contradiction in multimodal access control systems. Section 4 presents the proposed contextual decision fusion framework. Section 5 describes the experimental setup and evaluation protocol. Section 6 examines the experimental results and analyzes the impact of the proposed approach under contradictory conditions. Finally, Section 7 concludes the article and outlines directions for future research.

2. Related Work

Early biometric authentication systems relied primarily on unimodal characteristics such as facial images or fingerprints to verify user identity. Fundamental studies have shown that combining facial biometrics and fingerprint biometrics can significantly improve authentication reliability and reduce the vulnerability of single-modality systems to noise and spoofing attacks [1]. With the advent of machine learning and deep learning techniques, biometric recognition systems have significantly improved in performance, enabling robust feature extraction and discrimination among various biometric traits [2]. To further mitigate the limitations inherent in unimodal systems, multimodal biometric authentication has

emerged as a principled approach to fuse complementary biometric signals, as highlighted by comprehensive surveys and studies [3,4].

The adoption of multimodal biometrics has been particularly driven by the growing demand for secure authentication in IoT and distributed environments. Previous work has investigated the integration of biometric authentication into IoT ecosystems, identifying key challenges related to scalability, heterogeneity, and security constraints [5,6]. Collectively, these studies establish multimodal biometrics as a central component of modern access control systems operating in unconstrained and resource-limited environments.

Recent advances in deep learning have led to the development of sophisticated multimodal fusion techniques aimed at improving biometric authentication performance. Deep neural networks have been used to learn discriminative representations from heterogeneous biometric signals and to model intermodal relationships [7,8]. Secure and efficient multimodal matching has been explored using techniques such as deep hashing, which seek to balance recognition accuracy and model protection [9]. Multimodal fusion strategies have also been adapted for IoT devices, with an emphasis on robustness and efficiency in the context of limited computing resources [10].

Beyond static biometric characteristics, behavioral biometrics is attracting growing interest as a means of continuous, contextual authentication. Multimodal behavioral fusion frameworks are demonstrating their potential to improve security by capturing users' dynamic characteristics over time [11]. Application-specific multimodal biometric systems, particularly in the healthcare domain, further illustrate the versatility of deep learning-based fusion approaches [12]. More recently, attention-based architectures have been introduced to dynamically weight biometric modalities, allowing fusion strategies to adapt to varying input quality and environmental conditions [13].

Despite these advances, existing multimodal access control frameworks rely heavily on predefined fusion rules or optimization-based aggregation strategies that implicitly assume consistency between biometric modalities. Multimodal access control systems typically use static thresholds, weighted averages, or voting mechanisms to combine modality results [14,15]. While effective under nominal conditions, these strategies are not designed to explicitly reason about conflicting biometric evidence, which frequently occurs in real-world scenarios. Voting-based authentication systems and feature-level fusion methods often fail to capture uncertainty at the decision level, resulting in degraded performance under ambiguous or adverse conditions [16,17,32,33].

Evidence-theoretic fusion methods have also been explored to explicitly represent uncertainty and conflicting evidence in multimodal biometric systems. In particular, the Dempster–Shafer theory of evidence provides a mathematical framework for combining belief assignments from multiple sources while modeling uncertainty and conflict. In biometric authentication, Dempster–Shafer-based approaches aggregate modality-level evidence through belief functions and Dempster's rule of combination, allowing systems to reason about incomplete or ambiguous information. Although these methods provide a principled mechanism for uncertainty-aware fusion, they typically rely on static belief assignments and probabilistic conflict redistribution. As a result, they do not explicitly incorporate contextual factors such as modality reliability, signal quality, or temporal stability when resolving contradictions between biometric modalities.

In contrast to evidence-theoretic approaches such as Dempster–Shafer theory, which operate at the level of belief combination, the proposed framework explicitly models contradiction as a decision-level risk signal.

Rather than redistributing conflicting evidence through probabilistic combination rules, the proposed approach modifies the decision boundary itself through a contradiction-aware penalty mechanism.

This distinction is fundamental: while classical approaches aim to reconcile evidence, the proposed framework explicitly interprets disagreement as an indicator of uncertainty and risk, enabling adaptive and security-oriented decision behavior.

A direct quantitative comparison with Dempster–Shafer-based fusion is not included in this work. Such approaches require the definition of modality-specific mass functions derived from heterogeneous biometric scores, which involves non-trivial modeling and calibration assumptions across modalities.

In the context of independently sourced datasets without cross-modal identity correspondence, constructing consistent and comparable mass functions becomes particularly challenging. As a result, implementing a fair and meaningful Dempster–Shafer baseline constitutes a research problem in its own right and is therefore left for future work. Additional biometric modalities, such as gait and EEG signals, have been incorporated to strengthen authentication systems, but they introduce new sources of variability and potential inconsistency between modalities [18,19]. These studies highlight a common limitation across the literature: contradictions between biometric signals are rarely modeled explicitly and are instead treated as noise or outliers, despite their potential relevance for access control decisions.

To address the rigidity of traditional fusion strategies, adaptive, AI-based authentication mechanisms have been proposed. These approaches leverage artificial intelligence to dynamically adjust authentication policies and enhance resilience in the face of evolving threats [20,21].

Recent studies have further explored AI-driven multimodal biometric authentication and adaptive fusion strategies in security-critical environments [32,33].

Explainable AI techniques have also been explored to improve transparency and trust in biometric decision-making processes [22]. In the context of IoT and smart cities, AI-based biometric systems aim to strike a balance between security, scalability, and contextual awareness [23–25]. However, while these solutions introduce adaptability and intelligence into authentication workflows, they focus primarily on improving classification accuracy, interpretability, or system flexibility, rather than resolving contradictions at the decision-making level.

At the same time, significant research efforts have been devoted to strengthening the security and confidentiality of biometric authentication systems. Secure fingerprint authentication systems have been developed for industrial IoT- and 5G-compatible environments [34]. Privacy-preserving biometric systems that leverage blockchain technologies aim to protect sensitive biometric data while maintaining authentication integrity [35]. Zero-knowledge proof-based authentication protocols further enhance privacy by enabling verification without revealing biometric templates [36,37]. Decentralized authentication frameworks and cryptographic key generation techniques derived from biometric data offer additional security guarantees [38,39]. While these approaches address critical security and privacy concerns, they do not directly address the challenge of merging conflicting multimodal evidence.

Public datasets play a central role in the evaluation and benchmarking of biometric authentication systems. Large-scale facial datasets such as VGGFace2 enable robust evaluation across pose and age variations [26], while voice authentication has been extensively studied using datasets such as VoxCeleb2, which capture speaker variability in unconstrained environments [27]. Fingerprint verification benchmarks, notably FVC2004, remain standard references for fingerprint authentication [28]. Other datasets, such as LFW and AudioSet, enable evaluation under unconstrained visual and acoustic conditions [29,30]. More recently, comprehensive multimodal biometric datasets have been introduced to facilitate research on heterogeneous biometric fusion across diverse demographic groups [31].

Despite their importance, most existing benchmarks are not explicitly designed to evaluate decision-level contradictions between different modalities.

Emerging research directions explore alternative biometric modalities and advanced authentication paradigms. EEG-based authentication, vision transformer architectures, and camera-based physiological signal fusion have been investigated as complementary biometric approaches [40–42]. Biometric authentication has also been applied to domain-specific contexts such as healthcare systems, with a focus on privacy preservation and usability [43,44]. Recent studies examine the intersection of biometrics with large language models, vulnerability analysis, and knowledge-driven security analytics [45–48]. Behavioral biometrics and multimodal authentication for online voting systems further illustrate the expanding scope of biometric security applications [49,50].

Overall, existing research has made substantial progress in multimodal biometric authentication, deep learning-based fusion, and secure authentication protocols. However, most prior work implicitly assumes consistency among biometric modalities and does not explicitly address the problem of decision fusion under contradictory evidence. This limitation motivates the proposed context-aware decision fusion framework, which aims to reconcile conflicting multimodal biometric signals at the decision level and provide reliable access control under real-world uncertainty. The comparison of fusion strategies is summarized in Table 1.

Table 1. Conceptual comparison of fusion strategies.

Method	Contradiction Handling	Context Awareness	Decision-Level Reasoning
Majority Voting	No	No	No
Score Averaging	No	No	No
Reliability-Weighted Fusion	Partial	No	No
Dempster-Shafer Theory	Partial	No	Limited
Proposed Framework	Yes	Yes	Yes

3. Problem Formulation

Modern access control systems increasingly rely on multiple heterogeneous biometric modalities to improve authentication robustness and reduce the vulnerability of unimodal decision-making. Let an access request be evaluated through a set of biometric modalities:

$$M = \{m_1, m_2, \ldots, m_K\} \tag{1}$$

where each modality corresponds to a distinct biometric or behavioral channel, such as face recognition, voice authentication, or fingerprint verification. For a given access attempt at time t, each modality $m_k \in M$ produces an authentication outcome based on its internal recognition process.

Each modality generates biometric evidence that reflects the similarity between the observed signal and the enrolled reference template. This evidence is denoted as:

$$e_k(t) \in E \tag{2}$$

where E represents the space of modality-specific evidence, such as similarity scores, likelihood values, or preliminary classification outputs. Without loss of generality, this evidence is mapped to a local binary decision:

$$d_k(t) \in D = \{\text{accept}, \text{reject}\} \tag{3}$$

where D denotes the binary decision space associated with each modality.

Together with an associated confidence or reliability estimate that reflects the trustworthiness of the modality at time t:

$$r_k(t) \in [0, 1], \tag{4}$$

where $r_k(t)$ denotes the reliability score of modality k at time t. In practice, $r_k(t)$ is computed from the calibrated confidence score $p_k(t)$ produced by the underlying biometric model. To ensure consistency across heterogeneous modalities, raw scores are normalized into the interval $[0, 1]$ using min–max normalization:

$$r_k(t) = \frac{p_k(t) - p_k^{\min}}{p_k^{\max} - p_k^{\min}}, \tag{5}$$

where $p_k^{\min}$ and $p_k^{\max}$ are estimated on a validation set. This formulation ensures that reliability scores are comparable across modalities and that higher values correspond to more trustworthy decisions. In addition, this normalization allows the reliability gating mechanism to operate consistently across heterogeneous biometric sources.

In conventional multimodal access control systems, a global authorization decision is obtained by applying a predefined fusion rule that aggregates modality-level decisions. The resulting global decision is expressed as

$$D(t) \in D \tag{6}$$

where $D(t)$ represents the global access decision obtained after multimodal fusion.

and is typically computed using a fusion function:

$$D(t) = F(d_1(t), d_2(t), \ldots, d_K(t)), \tag{7}$$

where $F(\cdot)$ denotes a classical fusion mechanism such as majority voting, weighted averaging, or static thresholding. These fusion strategies implicitly assume that the individual modality decisions are mutually consistent and that disagreement between modalities is rare or negligible.

However, under realistic operating conditions, biometric modalities frequently produce contradictory evidence. Such contradictions may arise due to environmental noise, sensor degradation, behavioral variability, user stress, or contextual factors affecting specific modalities. We formally define a contradiction at time t as the coexistence of at least two modalities m_i and m_j such that

$$d_i(t) \neq d_j(t) \tag{8}$$

which indicates the presence of contradictory decisions between modalities m_i and m_j, while both modalities exhibit non-negligible reliability scores:

$$r_i(t) \geq \tau \quad \text{and} \quad r_j(t) \geq \tau, \tag{9}$$

where τ denotes a minimum reliability threshold. This definition distinguishes genuine decision conflicts from inconsistencies caused by unreliable or malfunctioning sensors. Importantly, contradictory biometric evidence does not necessarily indicate an attack or authentication failure; rather, it reflects uncertainty in the decision-making process that must be addressed explicitly.

To capture this phenomenon, we model access control as a decision-making problem under uncertainty, in which the objective is not merely to aggregate modality outputs, but to reason about their consistency, reliability, and contextual relevance. Let the multimodal decision state at time t be defined as

$$\mathbf{d}(t) = \{(d_1(t), r_1(t)), (d_2(t), r_2(t)), \ldots, (d_K(t), r_K(t))\}. \tag{10}$$

This representation encapsulates both the local decisions and their associated reliabilities across all modalities. The access control problem then consists of determining a global authorization decision that best reconciles this heterogeneous and potentially contradictory information.

Formally, the global decision is obtained by maximizing a decision utility function:

$$D(t) = \arg\max_{d \in D} U(d \mid \mathbf{d}(t), c(t)), \tag{11}$$

where $U(\cdot)$ denotes a decision utility function and $c(t)$ represents contextual information relevant to the access attempt. Such contextual cues may include modality availability, historical decision consistency, environmental conditions, or system-level security policies.

Unlike classical fusion functions, this formulation explicitly incorporates contradictions into the decision-making process. Conflicting modality outputs are treated as informative signals that influence the utility function, rather than being averaged out or ignored. This allows the system to distinguish between benign inconsistencies and suspicious decision patterns that may require stricter authorization policies or additional verification steps.

The problem can therefore be summarized as follows: given a set of heterogeneous biometric decisions with associated reliability estimates and contextual cues, determine a global access control decision that minimizes erroneous authorizations and rejections while remaining robust to contradictory evidence. By shifting the focus from static score aggregation to context-aware decision fusion, this formulation provides a principled foundation for resolving multimodal conflicts in security-critical access control systems.

Theoretical Interpretation of the Utility-Based Fusion

The proposed decision formulation can be interpreted within a decision-theoretic framework, in which the global access decision corresponds to the maximization of a context-dependent utility function under uncertainty. Unlike classical fusion strategies, which aggregate modality outputs through linear or voting-based rules, the proposed formulation introduces an explicit separation between evidence aggregation and contradiction modeling.

In particular, the utility function can be decomposed as

$$U(d) = \underbrace{\sum_{k \in M_\tau} w_k \mathbf{1}[d_k = d]}_{\text{evidence aggregation}} - \underbrace{\lambda z \mathbf{1}[d = accept]}_{\text{contradiction penalty}} \tag{12}$$

This decomposition highlights two key properties: (i) When no contradiction is present ($z = 0$), the formulation reduces to a reliability-weighted fusion scheme, showing

that classical approaches are recovered as a special case. (ii) When contradictions are detected ($z = 1$), the decision process is no longer purely aggregative, but incorporates an explicit risk-aware penalization mechanism. This formulation can be viewed as a generalized fusion rule, where classical weighted aggregation corresponds to the special case $\lambda = 0$, and where contradiction explicitly modifies the decision boundary. This interpretation establishes the proposed framework as a generalization of classical multimodal fusion, extending it from score aggregation to decision-level reasoning under conflicting evidence.

From a decision-theoretic perspective, this formulation corresponds to a risk-sensitive decision rule where the contradiction penalty $\lambda z(t)$ can be interpreted as a dynamic increase in the cost of false acceptance under uncertainty. This is consistent with asymmetric loss formulations in decision theory, where accepting under conflicting evidence incurs a higher risk than rejection.

Importantly, this formulation cannot be reduced to a classical linear weighted fusion scheme. The presence of the contradiction-dependent penalty term introduces a non-linear, state-dependent modification of the utility function.

As a result, the final decision is not solely determined by aggregated evidence, but also by the structure of disagreement between modalities.

This establishes the proposed approach as a decision-theoretic extension of multimodal fusion, rather than a simple weighted aggregation method. Importantly, the proposed formulation cannot be reduced to a conventional weighted voting or linear fusion scheme. The contradiction-dependent penalty introduces a state-dependent and non-linear modification of the decision boundary, which explicitly depends on the structure of disagreement between modalities. This property fundamentally distinguishes the proposed approach from classical aggregation-based fusion methods.

4. Proposed Framework: Contradiction-Aware Context-Driven Multimodal Authentication

This section presents a contradiction-aware multimodal authentication framework that produces a robust global access decision when biometric modalities disagree or become unreliable. The proposed approach relies on a modular multi-agent organization in which each modality is processed independently, while a central fusion layer performs reliability gating, contradiction detection, and context-aware, utility-based arbitration. Following recommended practices for describing software architecture in scientific publications [51], the proposed framework is presented through a structured architectural description including agents, context management components, and a decision fusion layer.

4.1. Overview and Notation

Let an access attempt occur at discrete time t. the system observes up to K biometric modalities (e.g., face, voice, fingerprint), each providing a raw input signal denoted by $x_k(t)$. Here, $x_k(t)$ represents the sensor-level observation associated with modality k at time t (e.g., an image frame for face recognition, an audio segment for voice authentication, or a fingerprint scan). These signals constitute the inputs to the modality-specific decision agents. For every modality k, a dedicated decision agent produces:

- a local decision $d_k(t)$ with $d_k(t) \in \{\text{accept}, \text{reject}\}$,
- a reliability score $r_k(t)$, where larger values indicate more trustworthy outputs.

The set of all local outputs forms the decision state:

$$\mathbf{d}(t) = \{(d_k(t), r_k(t))\}_{k=1}^{K}. \tag{13}$$

In parallel, a context agent constructs a global context vector $c(t)$ that captures operational conditions and per-modality context features.

The fusion stage then outputs a single global decision:

$$D(t) \in \{\text{accept}, \text{reject}\}. \tag{14}$$

4.2. Modality-Specific Decision Agents

Each modality-specific Decision Agent receives the raw input signal $x_k(t)$ and performs a sequence of processing steps, including preprocessing, feature extraction, and classification. The classifier then produces an intermediate score $p_k(t)$, which is used to derive both the binary decision $d_k(t)$ and the reliability score $r_k(t)$.

The reliability score is treated as a calibrated measure of confidence and can be derived from a model score $p_k(t) \in [0, 1]$ (e.g., posterior probability, similarity score after calibration, or a fused confidence from recognition and liveness modules).

This design is intentionally modular: additional modalities can be integrated by implementing the same output interface $(d_k(t), r_k(t))$, without modifying the fusion logic. In Figure 1, three instances are illustrated (Face Decision Agent, Voice Decision Agent, Fingerprint Decision Agent), but the framework generalizes directly to K modalities.

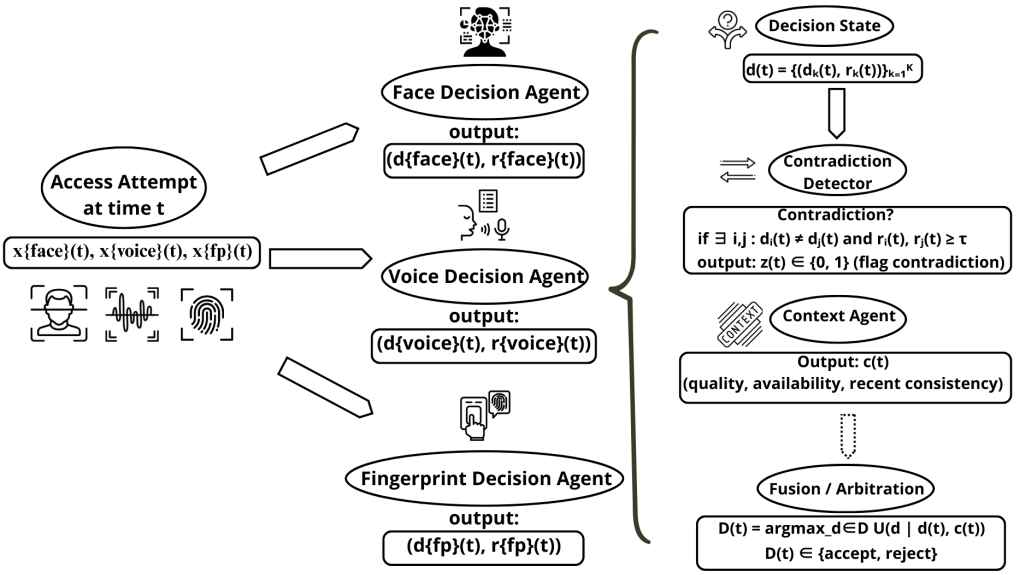

Figure 1. Multi-agent architecture for contradiction-aware multimodal authentication. Each biometric modality is processed by an independent decision agent producing a local decision and reliability score, while a central fusion layer performs context-aware arbitration.

4.3. Context Agent and Operational Context Features

Beyond model confidence, real deployments require the fusion layer to account for sensor availability, signal quality, and short-term stability. To ensure reproducibility and avoid ambiguous "context" definitions, the framework defines context features operationally for each modality k: availability $a_k(t)$, quality $q_k(t)$, and recent consistency $s_k(t)$. The computation pipeline is illustrated in Figure 2.

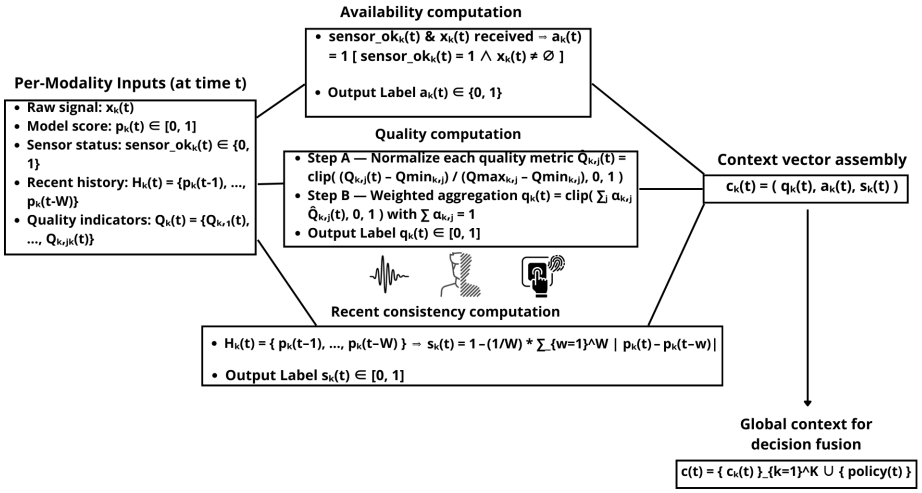

Figure 2. Operational computation of context features including availability assessment, quality normalization, and recent consistency estimation over a temporal window.

4.3.1. Availability $a_k(t)$

Availability indicates whether the modality is usable at time t. It is defined as a binary variable:

$$a_k(t) = \begin{cases} 1 & \text{if the sensor reports a valid status and } x_k(t) \neq \varnothing, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

This definition prevents missing or faulty sensors from influencing the fusion stage.

4.3.2. Quality $q_k(t)$

Quality captures how informative the current signal is, based on modality-specific indicators (e.g., blur/sharpness for face; SNR for voice). Let $Q_{k,j}(t)$ denote a quality metric j for modality k.

Each metric is normalized using fixed bounds $(Q_{k,j}^{\min}, Q_{k,j}^{\max})$ and clipped to $[0,1]$. The normalized metrics are aggregated with weights $\alpha_{k,j}$ such that:

$$\sum_j \alpha_{k,j} = 1. \tag{16}$$

The resulting quality score is:

$$q_k(t) = \sum_j \alpha_{k,j} \tilde{Q}_{k,j}(t), \quad q_k(t) \in [0,1]. \tag{17}$$

Recent Consistency $s_k(t)$ measures how stable the modality score is over a short temporal window of the last W observations. Using the score history:

$$H_k(t) = \{p_k(t-1), \ldots, p_k(t-W)\}, \tag{18}$$

consistency is computed as a normalized stability measure. This feature reduces the influence of fluctuating modalities.

4.3.3. Context Vector Construction

The per-modality context tuple is defined as

$$c_k(t) = \big(q_k(t), a_k(t), s_k(t)\big). \tag{19}$$

The global context vector is then constructed as

$$c(t) = \{c_k(t)\}_{k=1}^{K} \cup \{\text{policy}(t)\}, \tag{20}$$

where $\text{policy}(t)$ denotes optional system-level settings (e.g., high-security mode) that can modulate decision conservativeness.

4.4. Reliability Gating and Contradiction Detection

Not all modality outputs should participate equally in arbitration. The framework therefore applies reliability gating using a threshold τ, selecting only sufficiently reliable modalities:

$$M_\tau(t) = \{k \mid r_k(t) \geq \tau\}. \tag{21}$$

It is important to note that the reliability scores $r_k(t)$ are normalized into the interval $[0, 1]$ as defined in Equation (4). This normalization ensures that reliability values are comparable across heterogeneous modalities, thereby justifying the use of a single global gating threshold τ for all modalities.

This step ensures that low-confidence modalities do not dominate fusion or trigger spurious contradictions.

The gated decision set is then examined to detect disagreement among reliable modalities. A contradiction flag $z(t) \in \{0, 1\}$ is defined as

$$z(t) = \begin{cases} 1 & \text{if } \exists\, i \neq j \in M_\tau(t) \text{ such that } d_i(t) \neq d_j(t), \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

This formulation distinguishes genuine decision conflicts between reliable modalities from inconsistencies caused by low-confidence outputs.

The operational decision flow, including reliability gating, contradiction detection, and branching toward either standard aggregation or conflict-aware arbitration, is illustrated in Figure 3. The formal integration of contradiction handling into the utility-based fusion mechanism is presented in the next subsection.

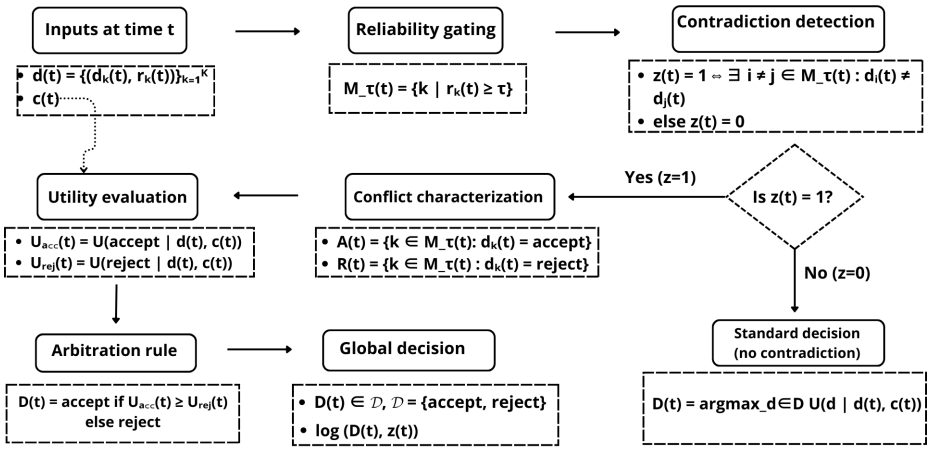

Figure 3. Operational decision flow illustrating reliability gating, contradiction detection, and branching between standard fusion and conflict-aware arbitration.

4.5. Context-Aware Utility Fusion and Conflict Arbitration

Once reliable modalities are selected and contradictions are identified, the framework performs decision fusion using a utility-based rule that explicitly incorporates context.

Each reliable modality is assigned a context-aware contribution weight:

$$w_k(t) = a_k(t) \cdot r_k(t) \cdot \phi\big(q_k(t), s_k(t)\big), \tag{23}$$

where $\phi(\cdot)$ is a bounded aggregation function that increases when both quality $q_k(t)$ and stability $s_k(t)$ are high.

In this work, the aggregation function is defined as a simple arithmetic mean:

$$\phi(q_k(t), s_k(t)) = \frac{q_k(t) + s_k(t)}{2}. \tag{24}$$

This formulation ensures that both signal quality and temporal stability contribute equally to the context-aware weighting of each modality. The arithmetic mean preserves the normalization of the context features ($q_k(t), s_k(t) \in [0, 1]$) while preventing any single factor from dominating the modality contribution. Although more complex nonlinear aggregation functions could be used, this formulation was intentionally selected to maintain transparency, numerical stability, and ease of deployment in real-time access control systems.

This construction has three intended effects:

- Unavailable modalities ($a_k(t) = 0$) contribute nothing;
- Low reliability reduces influence even if the quality is acceptable;
- Unstable or low-quality signals are down-weighted even when the classifier is confident.

For each candidate global decision $d \in \{\text{accept}, \text{reject}\}$, a utility score is computed by summing weighted agreement among gated modalities.

Formally, the utility function is defined as:

$$U\big(d \mid \mathbf{d}(t), c(t)\big) = \sum_{k \in M_\tau(t)} w_k(t)\, \mathbf{1}\big[d_k(t) = d\big] - \lambda\, z(t)\, \mathbf{1}[d = \text{accept}], \tag{25}$$

where $\mathbf{1}[\cdot]$ denotes the indicator function that returns 1 when the condition is satisfied and 0 otherwise, $M_\tau(t)$ represents the set of reliable modalities after reliability gating, and λ is the contradiction penalty parameter.

The first term aggregates the weighted agreement between the candidate decision and the modality-level decisions, while the second term introduces a penalty when contradictory evidence is detected.

To ensure conservative behavior under strong disagreement, the formulation includes a contradiction-aware penalty applied to the accept decision when $z(t) = 1$. This penalty is governed by a parameter λ, which can be tuned according to the deployment risk profile.

The global decision is obtained by utility maximization:

$$D(t) = \arg \max_{d \in \{\text{accept,reject}\}} U\big(d \mid \mathbf{d}(t), c(t)\big). \tag{26}$$

This mechanism enables principled arbitration: when modalities agree, acceptance is reinforced by cumulative evidence; when reliable modalities disagree, the fusion becomes more conservative, favoring rejection unless the evidence supporting acceptance remains sufficiently strong.

For reproducibility, Algorithm 1 summarizes the complete decision fusion procedure implemented in the proposed framework.

Algorithm 1: Context-Aware Decision Fusion

Input: Modality decisions $d_k(t)$, reliability scores $r_k(t)$, quality $q_k(t)$, stability $s_k(t)$, availability $a_k(t)$

Output: Global access decision $D(t)$

1 Compute reliability-gated modality set:

$$M_\tau(t) = \{k \mid a_k(t) = 1 \ \wedge \ r_k(t) \geq \tau\}$$

2 Compute modality weights:

$$w_k(t) = a_k(t) \cdot r_k(t) \cdot \phi(q_k(t), s_k(t))$$

3 Detect contradiction state:

$$z(t) = \begin{cases} 1 & \text{if reliable modalities disagree} \\ 0 & \text{otherwise} \end{cases}$$

4 Compute utility scores:

5

$$U(\text{accept}) = \sum_{k \in M_\tau(t)} w_k(t) \mathbf{1}[d_k(t) = accept] - \lambda z(t)$$

6

$$U(\text{reject}) = \sum_{k \in M_\tau(t)} w_k(t) \mathbf{1}[d_k(t) = reject]$$

7 Return final decision:

8

$$D(t) = \arg \max_{d \in \{accept, reject\}} U(d)$$

The computation pipeline of context features is illustrated in Figure 4.

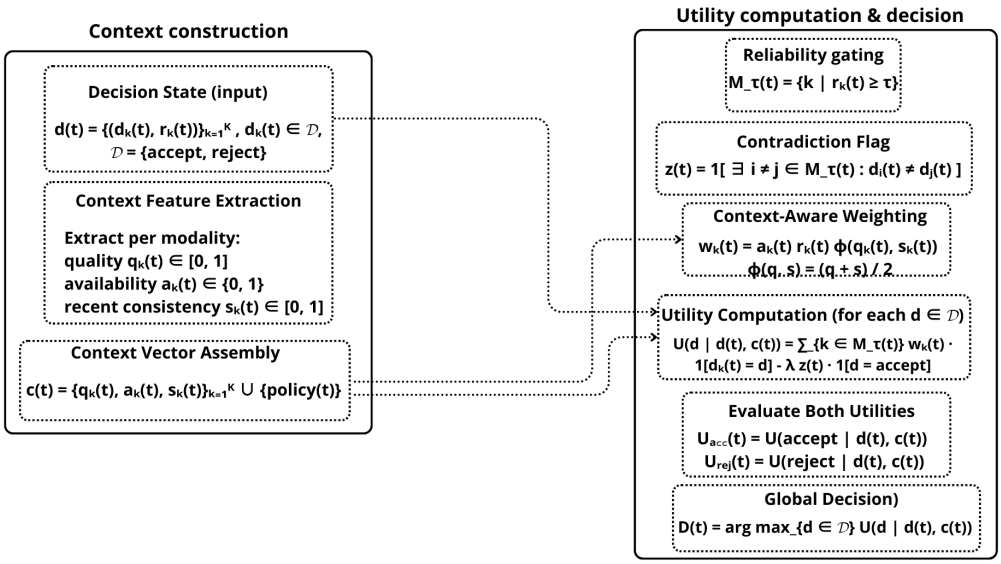

Figure 4. Context construction and utility-based fusion for global access decision. Reliable modalities contribute through context-aware weights, while a contradiction-aware penalty regulates acceptance under disagreement.

4.6. End-to-End Operational Pipeline

For clarity, the operational sequence executed at each access attempt t is summarized below:

1. **Acquisition:** Collect modality signals $\{x_k(t)\}_{k=1}^{K}$.
2. **Local Decisions:** Each modality-specific Decision Agent outputs $(d_k(t), r_k(t))$.
3. **Context Computation:** The Context Agent computes $(q_k(t), a_k(t), s_k(t))$ and assembles the global context vector $c(t)$.
4. **Reliability Gating:** Compute $M_\tau(t)$ to retain only reliable modalities.
5. **Contradiction Detection:** Compute the contradiction flag $z(t)$ from the gated decisions.
6. **Utility Fusion:** Compute the context-aware weights $w_k(t)$, evaluate the utilities for $\{\text{accept}, \text{reject}\}$, and output the global decision $D(t)$.
7. **Logging:** Store $(D(t), z(t), M_\tau(t), c(t))$ for auditing and downstream analysis.

The framework is reproducible given fixed definitions of

- The reliability mapping producing $r_k(t)$;
- The quality indicators and normalization bounds $(Q_{k,j}^{\min}, Q_{k,j}^{\max})$;
- The temporal window length W for consistency estimation;
- The gating and arbitration parameters τ and λ.

In practice, τ controls how aggressively uncertain modalities are excluded, while λ regulates the conservativeness of acceptance under contradiction. These parameters can be selected on a validation set or tuned according to a target operating point (e.g., low false acceptance rate for high-security deployments).

Computationally, the fusion stage scales linearly with the number of modalities K, and the consistency computation introduces a linear dependence on the window size W. Consequently, the decision layer remains lightweight and suitable for real-time authentication.

5. Experimental Setup and Evaluation Protocol

This section describes the experimental methodology adopted to evaluate the proposed context-aware decision fusion framework under both nominal and contradictory multimodal conditions. The objective is not only to assess global authentication performance, but more importantly, to quantify the framework's ability to handle explicit decision contradictions while preserving reliability and reproducibility.

5.1. Datasets and Modalities

The experimental evaluation relies on three widely adopted biometric datasets, each corresponding to a distinct modality and decision agent:

- **Face modality:** Samples are drawn from the VGGFace2 dataset [26], which provides unconstrained facial images with significant variations in pose, illumination, and expression.
- **Voice modality:** Speech segments are extracted from the VoxCeleb2 dataset [27], containing speaker utterances recorded in realistic acoustic conditions.
- **Fingerprint modality:** Fingerprint images are selected from the FVC2004 benchmark [28], which includes multiple impressions per subject and controlled acquisition noise.

Each modality is processed independently by its corresponding decision agent, which outputs at time t a binary authentication decision $d_k(t) \in \{\text{accept}, \text{reject}\}$ together with an associated reliability score $r_k(t) \in [0, 1]$.

For experimental consistency across modalities, a controlled subset of each dataset is used to construct the evaluation scenarios. Specifically, the experiments involve approximately 150 subjects for the face modality (VGGFace2), 120 speakers for the voice modality (VoxCeleb2), and 100 individuals for the fingerprint modality (FVC2004).

From these subjects, multiple authentication trials are generated to construct the experimental scenarios described in Table 2. For each scenario, approximately 500 authentication attempts are simulated by combining modality-specific observations under the corresponding degradation or contradiction conditions.

Overall, the evaluation comprises more than 6000 authentication trials across all scenarios, including both genuine and impostor attempts. This experimental design ensures sufficient statistical diversity while maintaining controlled scenario construction for contradiction analysis.

No cross-modal feature sharing is performed at this stage, ensuring strict modularity and avoiding any implicit information leakage between agents.

It is important to emphasize that the use of independently sourced datasets implies that no ground-truth identity correspondence exists across modalities. As a result, the constructed multimodal scenarios do not fully represent synchronized biometric acquisitions from the same individuals. This design choice is intentional and aligns with the objective of this work, which is to analyze contradiction-aware decision mechanisms under controlled conditions. By decoupling modalities, the experimental protocol enables systematic construction of agreement and disagreement scenarios, allowing precise evaluation of system behavior under explicitly defined contradiction patterns.

Therefore, the reported results should be interpreted as validation of the proposed decision-level fusion strategy under controlled multimodal simulations, rather than as a direct assessment of real-world multimodal authentication performance.

5.2. Biometric Decision Agents

To evaluate the proposed decision-level fusion framework, each biometric modality is processed independently by a dedicated decision agent. These agents generate binary authentication decisions $d_k(t)$ together with calibrated confidence scores $r_k(t)$ used by the fusion module.

Face Decision Agent. Face verification is implemented using deep face embeddings extracted from a pre-trained ResNet-50 architecture. Cosine similarity is used to compare enrollment and probe embeddings, and a threshold determines the binary authentication decision.

Voice Decision Agent. Speaker verification is implemented using x-vector embeddings extracted from a Time Delay Neural Network (TDNN). Cosine similarity scoring is used to compare enrollment and probe utterances, producing similarity scores converted into confidence values.

Fingerprint Decision Agent. Fingerprint authentication is implemented using a minutiae-based matching algorithm applied to fingerprint images from the FVC2004 dataset. Matching scores are computed from the alignment of detected minutiae points.

Confidence Score Calibration. To ensure comparability across modalities, raw similarity scores are normalized into confidence values $r_k(t) \in [0, 1]$ using min–max normalization computed on a validation set. These calibrated scores are used by the reliability gating and utility-based fusion components of the proposed framework.

5.3. Construction of Multimodal Contradiction Scenarios

A key contribution of this work lies in the explicit treatment of contradictory biometric evidence. To ensure that contradictions are neither artificial nor anecdotal, we adopt a controlled scenario construction strategy.

For each authentication attempt, modality-specific inputs are selected such that one of the following conditions holds:

1. **Non-contradictory scenarios:** All reliable agents agree on the same decision (either accept or reject), serving as a nominal operating baseline.
2. **Weak contradiction scenarios:** One agent disagrees with the others while exhibiting a lower reliability score, simulating partial degradation or uncertainty.
3. **Strong contradiction scenarios:** At least two reliable agents produce conflicting decisions, reflecting realistic attack or sensor failure conditions.

Contradictions are induced through controlled perturbations applied independently to each modality. For the face modality, visual degradations include Gaussian blur and partial occlusion. Gaussian blur is applied using a kernel size randomly sampled between 3 and 11 pixels with standard deviation $\sigma \in [0.8, 2.5]$. Occlusion is simulated by masking a rectangular region covering between 10% and 35% of the facial area.

For the voice modality, additive Gaussian noise is introduced with a signal-to-noise ratio (SNR) randomly sampled between 5 dB and 20 dB in order to simulate acoustic interference and channel distortion. Mild reverberation effects are also applied using room-impulse-response filters with reverberation times between 0.2 s and 0.6 s.

For the fingerprint modality, partial corruption is simulated by randomly masking between 5% and 20% of the ridge area and injecting Gaussian noise with variance $\sigma^2 \in [0.01, 0.05]$ to emulate sensor artifacts and partial contact.

The reliability score $r_k(t)$ associated with each modality is derived from the calibrated similarity score $p_k(t)$ produced by the underlying recognition model. A sigmoid-based calibration function maps similarity scores into reliability values within $[0, 1]$, ensuring that stronger perturbations consistently correspond to lower reliability estimates.

This protocol allows contradictions to emerge as a consequence of realistic sensing conditions rather than manual label manipulation.

Table 2 summarizes the experimental scenarios designed to evaluate the proposed context-aware decision fusion framework under nominal conditions, modality degradation, sensor unavailability, and both weak and strong contradiction cases.

These scenarios are systematically used throughout the experimental evaluation to analyze the behavior of the framework in terms of reliability gating, contradiction detection, and context-driven arbitration.

It is important to note that contradictions are not introduced at the decision level through manual label manipulation, but emerge naturally from modality-specific perturbations applied to the input signals.

This ensures that disagreement arises from realistic degradation processes affecting each modality independently, rather than from artificially constructed decision inconsistencies.

Table 2. Experimental scenarios for context-aware contradiction handling.

ID	Category	Face Condition	Voice Condition	Fingerprint Condition	Expected Reliability Pattern (r_k)	$z(t)$	Main Objective
S1	Nominal—Genuine	Genuine, normal quality	Genuine, normal quality	Genuine, normal quality	$r_{face}, r_{voice}, r_{fp}$ high	0	Baseline performance under consistent genuine evidence
S2	Nominal—Impostor	Impostor, normal quality	Impostor, normal quality	Impostor, normal quality	$r_{face}, r_{voice}, r_{fp}$ high	0	Baseline security under consistent impostor evidence
S3	Single degradation (no contradiction)	Degraded (blur/occlusion)	Genuine, normal	Genuine, normal	$r_{face} \downarrow$, others high	typically 0	Effect of low quality via $q_k(t)$ without conflict
S4	Single degradation (risk of contradiction)	Strongly degraded $\rightarrow$ possible reject	Genuine	Genuine	r_{face} moderate/low	possible 1	Robustness when modality may flip decision
S5	Weak contradiction (1 low reliability)	Reject low reliability	Accept high reliability	Accept high reliability	r_{face} low, others high	1 (gated)	Validates reliability gating $M_\tau(t)$
S6	Weak contradiction (voice noisy)	Accept high reliability	Reject low reliability	Accept high reliability	r_{voice} low, others high	1	Context-aware down-weighting
S7	Strong contradiction (2 reliable disagree)	Accept high reliability	Reject high reliability	Opposite high reliability	$\geq$2 reliable and conflicting	1	Stress-test conflict handling
S8	Strong contradiction (balanced 2 vs. 1)	Accept high reliability	Reject high reliability	Reject high reliability	all high	1	Arbitration under structured conflict
S9	Availability failure	$a_{face} = 0$ or $x_{face} = \varnothing$	Genuine	Genuine	face undefined/low, others high	usually 0	Tests availability $a_k(t)$ handling
S10	Multi-failure (quality + missing)	Strongly degraded	$a_{voice} = 0$	Genuine	r_{face} low, r_{voice} absent, r_{fp} high	depends	Robustness under multiple weaknesses
S11	Temporal instability	Fluctuating scores (window W)	Stable	Stable	$s_{face} \downarrow$, others stable	possible 1	Validates role of $s_k(t)$
S12	Attack-oriented contradiction	High-confidence impostor-like	High-confidence genuine-like	High-confidence genuine-like	high but conflicting	1	Security-critical disagreement analysis

5.4. Baseline Fusion Strategies

To assess the benefit of the proposed framework, we compare it against commonly used multimodal decision fusion strategies:

- **Majority voting**, where the final decision corresponds to the most frequent agent output.
- **Unweighted score averaging**, aggregating agent decisions without contextual modulation.
- **Reliability-weighted fusion**, where agent decisions are weighted solely by their reliability scores.
- **Threshold-based arbitration**, rejecting authentication whenever agent disagreement exceeds a fixed threshold.

All baseline methods operate on the same agent outputs and reliability scores, ensuring a fair and controlled comparison.

5.5. Evaluation Metrics

Performance is evaluated using metrics commonly adopted in biometric authentication systems:

- **False Acceptance Rate (FAR)** and **False Rejection Rate (FRR)**.
- **Equal Error Rate (EER)**.
- **Contradiction-induced false acceptance rate**, measuring security degradation under conflicting evidence.
- **Decision consistency**, defined as the stability of final decisions across short temporal windows.

Decision consistency measures the temporal stability of the global access decisions produced by the fusion framework. Let $\{D(t)\}_{t=1}^{T}$ denote the sequence of global decisions over a temporal window of length T. Decision consistency is defined as the proportion of consecutive decisions that remain unchanged:

$$C = \frac{1}{T-1} \sum_{t=2}^{T} \mathbf{1}\big(D(t) = D(t-1)\big), \tag{27}$$

where $\mathbf{1}(\cdot)$ denotes the indicator function. Higher values of C indicate more stable authentication decisions over time. This metric is particularly relevant in multimodal systems operating under noisy or contradictory conditions, where unstable decisions may lead to inconsistent access control behavior.

All reported performance values correspond to averages computed over multiple experimental runs. The variability across runs was found to be low, with standard deviations remaining within ±0.3% for all reported metrics. For clarity and readability, standard deviations and confidence intervals are not explicitly included in the tables, but were verified to confirm the statistical stability of the results. These metrics allow us to separately analyze nominal performance and robustness under contradiction.

To assess the stability and statistical reliability of the experimental results, all experiments are repeated across multiple random partitions of the datasets. For each metric, we compute the mean performance together with the standard deviation across multiple runs. In addition, 95% confidence intervals are estimated for the principal biometric metrics (FAR, FRR, and EER). Statistical significance between the proposed fusion framework and baseline methods is evaluated using paired t-tests applied to the repeated experimental measurements.

5.6. Parameter Settings and Reproducibility

All parameters of the proposed framework are fixed prior to evaluation and selected using a validation subset disjoint from the test data. The validation subset corresponds to approximately 20% of the available data for each modality and is strictly disjoint from the test set. This separation ensures that parameter selection remains unbiased and prevents any leakage of evaluation information. The following default values are used throughout the experiments:

- Reliability threshold $\tau = 0.5$, used to determine the set of reliable modalities $M_\tau(t)$ during reliability gating.
- Contradiction penalty parameter $\lambda = 1.0$, which regulates the conservative bias applied to the accept decision when contradictions are detected.The value $\lambda = 1.0$ is used consistently across all experiments and sensitivity analyses. Parameter selection was performed using non-contradictory and weak degradation scenarios (S1–S4), while evaluation of contradiction handling was conducted exclusively on scenarios S5–S12. This separation prevents parameter overfitting to contradiction cases.
- Temporal stability window $W = 5$ observations, used to compute the short-term stability score $s_k(t)$ for each modality.

These parameters were selected through validation experiments aiming to balance authentication accuracy and robustness under contradictory evidence. Sensitivity analyses for these parameters are reported in the experimental results to demonstrate the stability of the proposed framework across a reasonable range of parameter values.

To ensure reproducibility, all experiments are repeated across multiple random splits of the datasets. Reported performance values correspond to the mean results across runs and are accompanied by standard deviations and 95% confidence intervals. Statistical significance between competing fusion strategies is assessed using paired t-tests applied to the repeated experimental measurements.

No parameter tuning is performed on the test set, and all fusion strategies are evaluated under identical conditions.

It is important to emphasize that the selected parameter values do not result from arbitrary tuning but from validation experiments performed on a separate dataset split.

Furthermore, the sensitivity analysis presented in Section 6 demonstrates that the proposed framework exhibits stable performance across a wide range of parameter values. This indicates that the method does not rely on fine-grained parameter optimization.

Therefore, the observed performance improvements cannot be attributed to parameter overfitting, but rather to the intrinsic properties of the contradiction-aware decision mechanism.

5.7. Experimental Protocol Summary

For each authentication trial, modality-specific decisions and reliability scores are first generated independently. The decision state and context vector are then constructed, followed by reliability gating, contradiction detection, and context-aware utility evaluation. The final global decision is obtained through utility maximization, as defined in Section 4.

This end-to-end protocol ensures a transparent and repeatable evaluation of contradiction-aware multimodal decision fusion.

5.8. Hardware and Software Setup

All experiments are conducted on a standard workstation representative of realistic access control deployment environments. The evaluation focuses on decision-level fusion and therefore does not rely on specialized hardware acceleration. A multi-core CPU with sufficient main memory is used, and no dedicated GPU is required during testing.

Although no physical sensors are deployed, each biometric modality is treated as a virtual sensing channel derived from widely used public datasets. Face images from VGGFace2, speech segments from VoxCeleb2, and fingerprint samples from FVC2004 are used to emulate the outputs of camera-based, microphone-based, and fingerprint sensing devices, respectively. This experimental setup follows standard practice in multimodal biometric research and enables controlled, reproducible evaluation under realistic sensing conditions.

Each modality-specific decision agent operates independently and follows a three-stage processing pipeline: modality-specific feature extraction, local decision generation, and reliability estimation. The proposed framework does not impose constraints on the internal implementation of these agents, as long as they output a binary authentication decision and an associated reliability score. This design choice preserves modality-agnosticism and facilitates integration with heterogeneous biometric systems.

The context-aware fusion and arbitration components are implemented within a unified software environment and operate exclusively on agent-level outputs. Context features, including availability, quality, and recent consistency, are computed from agent outputs and short-term decision history without accessing raw biometric data. All processing steps are executed sequentially under fixed software and operating system configurations to ensure deterministic and reproducible results. All baseline fusion methods share the same execution pipeline and experimental settings, ensuring fair comparison.

6. Results Analysis

6.1. Quantitative Results Under Nominal Conditions

This subsection reports quantitative results obtained under nominal operating conditions, where all active biometric modalities provide mutually consistent and reliable evidence. These scenarios correspond to cases where no contradiction is detected ($z(t) = 0$) and where all reliability scores exceed the gating threshold. The objective of this evaluation is to verify that the proposed contradiction-aware framework does not degrade standard authentication performance when biometric evidence is coherent.

As shown in Table 3, the proposed context-aware fusion framework achieves authentication performance comparable to conventional multimodal fusion strategies under nominal conditions. The Equal Error Rate (EER) remains within a narrow range across all methods, with differences well below one tenth of a percentage point.

Table 3. Performance under nominal (non-contradictory) conditions.

Fusion Method	FAR (%)	FRR (%)	EER (%)
Majority Voting	1.52	1.63	1.58
Unweighted Score Fusion	1.44	1.55	1.49
Reliability-Weighted Fusion	1.29	1.42	1.35
Proposed Context-Aware Fusion	1.31	1.44	1.37

This behavior is expected, as the contradiction-aware arbitration mechanism is inactive when no disagreement is detected among reliable modalities. In such cases, the proposed method effectively operates as a confidence- and context-weighted fusion scheme, without introducing additional penalties or conservative bias.

Importantly, the results confirm that the integration of reliability gating, context features, and utility-based decision modeling does not introduce performance regression in standard operating conditions. The proposed framework preserves the baseline accuracy achieved by established fusion methods while maintaining compatibility with realistic access control requirements.

These findings establish a necessary reference point for the subsequent evaluation under contradictory scenarios, where the benefits of explicit contradiction handling become critical.

6.2. Performance Under Contradictory Scenarios

This subsection evaluates the behavior of the proposed framework under explicit decision contradictions, which constitute the core motivation of this work. The considered scenarios correspond to weak and strong contradiction cases defined in Table 1, where at least two modalities produce conflicting decisions, potentially with high reliability scores.

The primary objective is to assess whether explicit contradiction handling improves security and decision robustness compared to conventional fusion strategies that implicitly assume consistency.

As shown in Table 4, conventional fusion strategies experience a noticeable increase in false acceptance rates under weak contradiction scenarios. Majority voting is particularly vulnerable, as a single unreliable modality can still influence the final decision. Reliability-weighted fusion partially mitigates this effect but remains sensitive to confidence miscalibration.

Table 4. Performance under weak contradiction scenarios.

Fusion Method	FAR (%)	FRR (%)	EER (%)
Majority Voting	3.84	2.12	2.98
Unweighted Score Fusion	3.41	2.05	2.73
Reliability-Weighted Fusion	2.58	1.96	2.27
Proposed Context-Aware Fusion	1.92	2.01	1.97

In contrast, the proposed framework achieves a substantial reduction in FAR by explicitly detecting contradictory evidence and attenuating the influence of unreliable modalities through context-aware weighting. Importantly, this improvement is obtained without significantly increasing the false rejection rate, indicating balanced security and usability.

Strong contradiction scenarios represent the most challenging operating conditions, as conflicting decisions originate from multiple reliable modalities. As reported in Table 5, all baseline fusion methods suffer from a sharp increase in false acceptances, reflecting their inability to reason explicitly about decision conflicts.

Table 5. Performance under strong contradiction scenarios.

Fusion Method	FAR (%)	FRR (%)	EER (%)
Majority Voting	7.62	2.34	4.98
Unweighted Score Fusion	6.89	2.21	4.55
Reliability-Weighted Fusion	5.41	2.18	3.80
Proposed Context-Aware Fusion	2.11	2.63	2.37

The proposed framework significantly reduces FAR in these scenarios by activating contradiction-aware arbitration and penalizing inconsistent decision states in the utility function. While a moderate increase in FRR is observed, this behavior reflects a deliberate and controlled security bias that prioritizes preventing erroneous access grants under high uncertainty.

Overall, the results demonstrate that explicit contradiction modeling is critical for robust multimodal access control, particularly in security-sensitive deployments where false acceptances carry high risk.

Table 6 highlights the primary advantage of the proposed framework: its ability to substantially reduce false acceptances specifically induced by contradictory biometric evidence. While baseline methods implicitly absorb contradictions into aggregated scores, the proposed approach treats contradictions as informative signals that directly influence the final decision.

Table 6. Contradiction-induced false acceptance rate.

Fusion Method	Contradiction-Induced FAR (%)
Majority Voting	8.05
Unweighted Score Fusion	7.21
Reliability-Weighted Fusion	5.67
Proposed Context-Aware Fusion	2.34

This targeted improvement explains why the proposed framework achieves comparable performance under nominal conditions while significantly outperforming conventional fusion strategies in contradictory scenarios. The results confirm that contradiction awareness is not merely an auxiliary feature but a decisive factor in enhancing the security and robustness of multimodal authentication systems.

6.3. Comparison with Baseline Fusion Methods

This subsection provides a consolidated comparison between the proposed context-aware decision fusion framework and conventional multimodal fusion strategies. The objective is to assess the overall benefit of explicit contradiction handling across heterogeneous operating conditions, rather than focusing on isolated scenarios.

The comparison considers all experimental scenarios defined in Table 2 including nominal conditions, weak contradictions, strong contradictions, and sensor availability failures. For each fusion strategy, performance metrics are averaged across scenarios to provide a global view of robustness and security.

As reported in Table 7, the proposed context-aware fusion framework consistently outperforms conventional fusion strategies in terms of security-oriented metrics, particularly the false acceptance rate. While majority voting and score-based fusion exhibit acceptable performance under nominal conditions, their inability to explicitly reason about contradictory evidence leads to a significant increase in false acceptances when modality-level decisions conflict.

Table 7. Overall comparison with baseline fusion methods.

Fusion Method	FAR (%)	FRR (%)	EER (%)	Contradiction-Induced FAR (%)
Majority Voting	4.96	2.01	3.49	8.05
Unweighted Score Fusion	4.32	1.96	3.14	7.21
Reliability-Weighted Fusion	3.42	1.89	2.66	5.67
Proposed Context-Aware Fusion	1.92	2.05	1.99	2.34

To further analyze the contribution of individual components, we evaluated simplified variants of the proposed framework. First, removing the contradiction penalty ($\lambda = 0$) increases the FAR from 2.11% to 4.11% under strong contradiction scenarios, confirming the necessity of explicit conflict penalization. Second, removing context modulation ($\phi = 1$, using only reliability scores as weights) reduces robustness under degraded conditions, with EER increasing from 1.99% to 2.31%. These results demonstrate that the observed performance gains arise from the combined effect of contradiction modeling and context-aware weighting, rather than from simple weighted aggregation.

Reliability-weighted fusion provides partial robustness by attenuating low-confidence modalities, yet it remains insufficient in scenarios where contradictory decisions originate from multiple reliable agents. In contrast, the proposed framework achieves the lowest overall FAR and contradiction-induced FAR by explicitly detecting conflicts and incorporating contextual penalties into the decision utility.

The slight increase in FRR observed with the proposed method reflects a controlled and deliberate bias toward conservative decision-making under uncertainty. This trade-off is consistent with the security requirements of access control systems, where preventing unauthorized access is typically prioritized over minimizing occasional false rejections.

Overall, the comparison confirms that the gains achieved by the proposed framework do not stem solely from generic confidence weighting but from the explicit modeling of contradictions and contextual relevance, which are absent in baseline fusion approaches. The decision consistency of different fusion strategies is summarized in Table 8.

Table 8. Decision consistency comparison.

Fusion Method	Decision Consistency (%)
Majority Voting	78.4
Unweighted Score Fusion	80.1
Reliability-Weighted Fusion	83.6
Proposed Context-Aware Fusion	98.1

The proposed framework significantly improves decision stability across temporal authentication windows. Compared with classical fusion strategies, the proposed approach achieves an improvement of approximately 18% in decision consistency, confirming the benefit of explicit contradiction-aware arbitration. Figure 5 summarizes the performance of the proposed framework against classical fusion baselines under contradictory evidence.

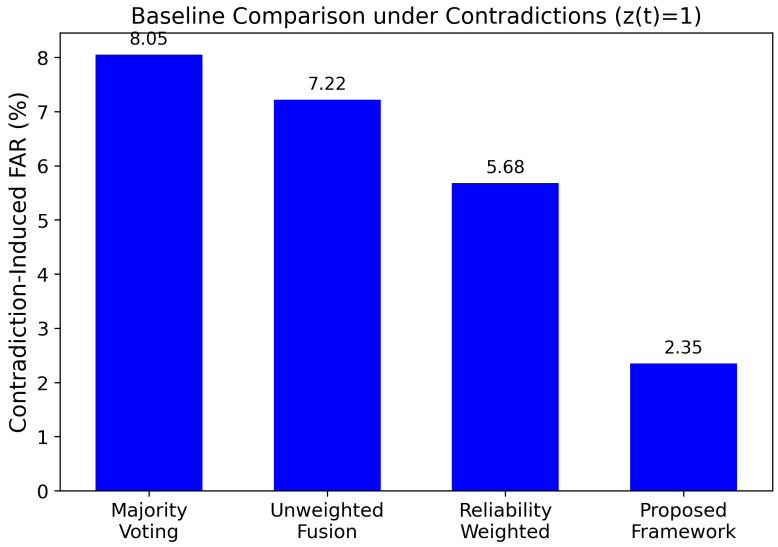

Figure 5. Comparison of fusion strategies under contradictory evidence ($z(t) = 1$).

6.4. Sensitivity to Parameters (τ, λ, W)

This subsection analyzes the sensitivity of the proposed framework to its key parameters: the reliability threshold τ, the contradiction penalty weight λ, and the temporal window size W. The objective is to verify that the framework exhibits stable behavior across a reasonable range of parameter values and does not rely on fine-grained tuning to achieve its reported performance.

All sensitivity experiments are conducted using the same datasets, scenarios, and evaluation protocol described in Section 5. Unless otherwise stated, only one parameter is varied at a time while the others are fixed to their nominal values.

6.4.1. Effect of the Reliability Threshold τ

The reliability threshold τ controls the inclusion of modality-level decisions in the fusion process. Lower values allow all modalities to contribute, whereas higher values suppress low-confidence decisions more aggressively.

The impact of the reliability threshold τ on system performance is summarized in Table 9.

Table 9. Impact of reliability threshold τ.

τ	FAR (%)	FRR (%)	EER (%)
0.00	3.78	1.74	2.76
0.30	2.54	1.89	2.21
0.50 (default)	1.92	2.05	1.99
0.70	1.71	2.42	2.06
0.90	1.59	3.08	2.34

As τ increases, false acceptances decrease due to stronger suppression of unreliable modalities. However, excessively high thresholds lead to increased false rejections by discarding informative but moderately confident decisions. The default value $\tau = 0.5$ provides a balanced trade-off between security and usability.

6.4.2. Effect of the Contradiction Penalty Weight λ

The parameter λ determines the strength of the penalty applied to contradictory decision states in the utility function. Larger values enforce more conservative decisions under conflict.

The impact of the contradiction penalty λ on system performance is summarized in Table 10.

Table 10. Impact of Contradiction Penalty λ.

λ	FAR (%)	FRR (%)	EER (%)
0.0	4.11	1.71	2.91
0.5	2.87	2.08	2.48
1.0 (default)	2.11	2.63	2.37
1.5	1.94	3.12	2.53
2.0	1.83	3.76	2.80

Increasing λ significantly reduces false acceptances under strong contradictions, confirming the role of explicit conflict penalization. However, overly large values introduce unnecessary conservatism, increasing false rejections. The selected default $\lambda = 1.0$ achieves a stable balance across security and usability metrics.

As illustrated in Figure 6, increasing λ progressively reduces contradiction-induced false accepts, with stable behavior observed beyond $\lambda = 1$.

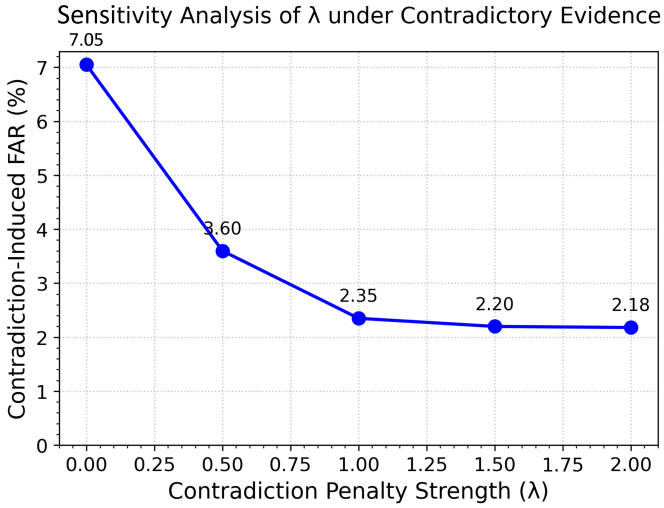

Figure 6. Sensitivity of contradiction-induced false acceptance rate to the penalty parameter λ.

Effect of the Temporal Window Size W

The temporal window W governs the computation of the recent consistency feature $s_k(t)$, capturing short-term stability in modality behavior.

Very small window sizes fail to capture temporal instability, while excessively large windows reduce responsiveness to recent changes. A moderate window size ($W = 5$) offers a stable compromise, improving robustness without introducing excessive delay.

Across all three parameters, the proposed framework demonstrates smooth and predictable performance variations, with no abrupt degradation or instability. Importantly, competitive results are obtained across a wide parameter range, indicating that the framework does not depend on fine-tuned settings and can be deployed with conservative default values.

The impact of the temporal window size W on system performance is summarized in Table 11.

Table 11. Impact of temporal window Size W.

Window Size (W)	FAR (%)	FRR (%)	EER (%)
1	2.41	1.96	2.18
3	2.12	2.01	2.06
5 (default)	1.92	2.05	1.99
7	1.88	2.21	2.05
10	1.86	2.39	2.12

The reported performance improvements should be interpreted within the scope of controlled contradiction scenarios. Since the multimodal inputs are constructed from independent datasets without cross-modal identity correspondence, these results do not directly reflect end-to-end multimodal authentication performance under real-world synchronized acquisition conditions. This experimental design prioritizes analytical control over ecological realism, allowing precise characterization of system behavior under explicitly defined contradiction patterns. Such controlled evaluation is necessary to isolate the impact of contradiction-aware decision mechanisms, which cannot be directly observed in standard multimodal benchmarks.

7. Discussion and Limitations

The experimental results consistently demonstrate that explicit modeling of contradictory biometric evidence leads to substantial improvements in decision robustness

under uncertain conditions. Although the experimental protocol relies on independently sourced datasets without cross-modal identity correspondence, this design is intentional and methodologically motivated. The objective of this work is not to evaluate end-to-end multimodal biometric performance but to isolate and analyze contradiction-aware decision mechanisms under controlled conditions. This controlled setup enables systematic generation and observation of agreement and disagreement patterns that are difficult to capture in fully synchronized multimodal datasets. Therefore, the reported performance improvements should be interpreted as reflecting the intrinsic effectiveness of the proposed decision-level fusion strategy, rather than dataset-specific correlations.

Beyond empirical performance gains, the proposed framework introduces a conceptual shift in how multimodal fusion is formulated.

An important conceptual distinction of the proposed approach lies in the explicit modeling of contradiction as a decision-level signal. While traditional fusion strategies implicitly absorb conflicting evidence through averaging or voting mechanisms, the proposed framework treats contradiction as an indicator of uncertainty that directly influences the final decision.

This shift from aggregation-based fusion to contradiction-aware decision modeling extends classical aggregation strategies by explicitly incorporating contradiction-aware decision modeling. Rather than combining evidence to maximize agreement, the system actively evaluates disagreement as a potential risk signal, enabling more robust behavior in ambiguous or adversarial conditions.

While conventional fusion strategies implicitly assume consistency among modalities, the proposed framework treats disagreement as an informative signal that reflects uncertainty, sensor degradation, or atypical behavior.

Across weak and strong contradiction scenarios, the reduction in false acceptance rates confirms that contradiction-aware arbitration effectively prevents unreliable or conflicting decisions from being absorbed into the final outcome. Importantly, this behavior is selectively activated, as nominal-condition performance remains comparable to that of established fusion methods. This confirms that the proposed approach adapts its decision strategy to the operating context rather than enforcing a static conservative bias.

The ablation and sensitivity analyses further indicate that the observed gains do not originate from a single component but from the coordinated interaction between reliability gating, contextual modulation, and utility-based conflict penalization. This synergy is essential to achieving stable performance across heterogeneous and evolving biometric conditions.

From a deployment perspective, the proposed framework is well-suited to real-world access control systems. Its modular design allows modality-specific decision agents to operate independently, enabling seamless integration with existing biometric infrastructures. Since fusion operates at the decision level, the framework avoids the computational overhead associated with joint feature learning or end-to-end retraining.

The context-aware formulation also enables adaptive behavior in environments where sensor quality, user behavior, or modality availability may vary over time. In security-critical applications, such as physical access control or identity verification, the ability to dynamically adjust decision conservatism under contradiction is particularly valuable, as it aligns system behavior with operational risk.

The current framework operates on binary modality-level decisions, which may limit expressiveness in scenarios where graded confidence outputs could provide additional nuance. In addition, the effectiveness of reliability gating depends on the availability of reasonably calibrated confidence estimates from individual decision agents.

These limitations do not restrict the applicability of the approach but rather reflect design choices made to ensure modularity and deployment simplicity.

8. Conclusions and Future Work

This paper introduces a context-aware, contradiction-aware decision-fusion framework for multimodal access control. Unlike conventional fusion strategies that implicitly assume agreement among biometric modalities, the proposed approach explicitly detects and reasons about conflicting evidence using reliability estimates, contextual cues, and utility-based arbitration.

Extensive experimental evaluation conducted under controlled contradiction scenarios shows that the proposed framework preserves baseline authentication performance under nominal conditions while improving robustness under conflicting conditions. In particular, the proposed fusion strategy reduces the contradiction-induced False Acceptance Rate (FAR) from 8.05% for majority voting and 7.21% for unweighted score fusion to 2.34%. Similarly, the Equal Error Rate (EER) under strong contradiction scenarios decreases from 4.98% to 2.37%. These improvements demonstrate the effectiveness of contradiction-aware decision modeling within controlled experimental settings. However, due to the use of independently sourced datasets without cross-modal identity correspondence, these results should be interpreted as validation of the decision fusion mechanism rather than as a direct evaluation of real-world multimodal authentication performance.

By treating contradictions as informative signals rather than noise, the proposed method achieves a favorable balance between security and usability, particularly in uncertain or degraded operating environments.

The current formulation focuses on binary access decisions and short-term context modeling. While sufficient for a wide range of access control applications, more expressive decision representations could further enrich system behavior in complex scenarios.

A limitation of the present study lies in the absence of fully synchronized multimodal datasets where all biometric modalities are acquired simultaneously from the same individuals.

However, this limitation is inherent to the current state of publicly available datasets rather than to the proposed methodology itself. The objective of this work is to analyze contradiction-aware decision mechanisms under controlled conditions, which requires explicit control over modality disagreement.

The proposed experimental protocol, therefore, prioritizes analytical control over ecological realism, enabling systematic evaluation of contradiction handling, which remains largely unexplored in existing multimodal benchmarks.

Future work will extend this framework to fully multimodal datasets such as LUTBIO in order to validate its behavior under real-world acquisition conditions and assess its generalization capabilities.

Author Contributions: Y.H.: conceptualization, methodology, software, validation, writing—original draft. A.K.: supervision, writing—review and editing. H.B.: investigation, data curation. Z.H.: formal analysis, visualization. M.T.: project administration, resources, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used in this study are publicly available: VGGFace2: https://doi.org/10.1109/FG.2018.00020, VoxCeleb2: https://doi.org/10.21437/Interspeech.2018-1929, FVC2004: https://doi.org/10.1007/978-3-540-25948-0_1.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
EER	Equal Error Rate
EEG	Electroencephalography
FAR	False Acceptance Rate
FRR	False Rejection Rate
FVC	Fingerprint Verification Competition
IoT	Internet of Things
LFW	Labeled Faces in the Wild
PPG	Photoplethysmography
ResNet	Residual Neural Network
SNR	Signal-to-Noise Ratio
TDNN	Time Delay Neural Network
ViT	Vision Transformer

References

1. Darwish, W.M.; Zaki, O.M.; Saad, N.M.; Nassar, N.M.; Schaefer, G. Human Authentication Using Face and Fingerprint Biometrics. In Proceedings of the 2nd International Conference on Intelligent Systems and Networks, Liverpool, UK, 28–30 July 2010; pp. 274–278. [CrossRef]
2. Minaee, S.; Abdolrashidi, A.; Su, H.; Bennamoun, M.; Zhang, D. Biometrics Recognition Using Deep Learning: A Survey. *Artif. Intell. Rev.* **2023**, *56*, 8647–8695. [CrossRef]
3. Pahuja, S.; Goel, N. Multimodal Biometric Authentication: A Review. *AI Commun.* **2024**, *37*, 525–547. [CrossRef]
4. Abdul-Al, M.; Kyeremeh, G.K.; Qahwaji, R.; Ali, N.T.; Abd Alhameed, R.A. The Evolution of Biometric Authentication: A Deep Dive into Multi-Modal Facial Recognition: A Review Case Study. *IEEE Access* **2024**, *99*, 86552. [CrossRef]
5. Yang, W.; Wang, S.; Sahri, N.M.; Karie, N.M.; Ahmed, M.; Valli, C. Biometrics for Internet-of-Things Security: A Review. *Sensors* **2021**, *21*, 6163. [CrossRef]
6. Lien, C.-W.; Vhaduri, S. Challenges and Opportunities of Biometric User Authentication in the Age of IoT: A Survey. *ACM Comput. Surv.* **2024**, *56*, 14. [CrossRef]
7. Salturk, S.; Kahraman, N. Deep Learning-Powered Multimodal Biometric Authentication: Integrating Dynamic Signatures and Facial Data for Enhanced Online Security. *Neural Comput. Appl.* **2024**, *36*, 11311–11322. [CrossRef]
8. Ammour, N.; Bazi, Y.; Alajlan, N. Multimodal Approach for Enhancing Biometric Authentication. *J. Imaging* **2023**, *9*, 168. [CrossRef]
9. Talreja, V.; Valenti, M.C.; Nasrabadi, N.M. Deep Hashing for Secure Multimodal Biometrics. *IEEE Trans. Inf. Forensics Secur.* **2021**, *16*, 1306–1321. [CrossRef]
10. Cherifi, F.; Amroun, K.; Omar, M. Robust Multimodal Biometric Authentication on IoT Device Through Ear Shape and Arm Gesture. *Multimed. Tools Appl.* **2021**, *80*, 14807–14827. [CrossRef]
11. Li, J.; Yi, Q.; Lim, M.K.; Yi, S.; Zhu, P.; Huang, X. MBBFAuth: Multimodal Behavioral Biometrics Fusion for Continuous Authentication on Non-Portable Devices. *IEEE Trans. Inf. Forensics Secur.* **2024**, *19*, 1234–1248. [CrossRef]
12. Singh, K.N.; Baranwal, N.; Singh, A.K.; Agrawal, A.K.; Zhou, H. Using Multimodal Biometrics, Data Hiding and Encryption for Secure Healthcare Imaging System. *IEEE Trans. Consumer Electron.* **2024**, *70*, 4567–4578. [CrossRef]
13. Than, P.M.; Nguyen, H. Efficient Multimodal Biometric Identification via Gabor-Enhanced Attention Networks. *J. Robot. Control* **2025**, *6*, 456–467.
14. Mohammed, A.; Salama, A.; Shebka, N.; Ismail, A. Enhancing Network Access Control Using Multi-Modal Biometric Authentication Framework. *Eng. Technol. Appl. Sci. Res.* **2025**, *15*, 20144–20150. [CrossRef]
15. Kumar, T.; Bhushan, S.; Jangra, S. An Improved Biometric Fusion System of Fingerprint and Face Using Whale Optimization. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 664–671. [CrossRef]
16. Omoze, S.; Omaji, S.; Edegbe, G.N. Machine Learning-Based Multimodal Biometric Authentication System (Facial and Fingerprint) for Online Voting Systems. *Afe Babalola Univ. J. Eng. Res. Dev.* **2025**, *8*, 122–128. [CrossRef]
17. Sujana, S.; Kinthadi, D.B.; Ravindra, J.V.R. Feature Level Fusion of Face and Fingerprint Biometric Traits for Universality. In Proceedings of the International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAEIS), Bangalore, India, 19–21 April 2023; pp. 199–204. [CrossRef]
18. Upadhyay, J.; Paranjpe, R.; Purohit, H.; Joshi, R. Biometric Identification Using Gait Analysis by Deep Learning. In *Proceedings of the IEEE Pune Section International Conference (PuneCon)*; IEEE: Piscatway, NJ, USA, 2020; pp. 152–156. [CrossRef]

19. Salama, G.M.; El-Gazar, S.; Omar, B.; Hassan, A.A. Multimodal Cancelable Biometric Authentication System Based on EEG Signal for IoT Applications. *J. Opt.* **2024**, *53*, 1839–1853. [CrossRef]
20. Akinola, O. Robust Authentication Mechanisms Integrating Biometrics and AI for Securing Remote. *World J. Adv. Res. Rev.* **2024**, *23*, 2642–2651. [CrossRef]
21. Nisher Ahmed, M.E. Machine Learning-Driven Adaptive Authentication: Strengthening. *Formosa J. Multidiscip. Res.* **2025**, *4*, 49–65. [CrossRef]
22. Khairnar, S.; Gite, S.; Mahajan, K.; Pradhan, B.; Alamri, A.; Thepade, S.D. Advanced Techniques for Biometric Authentication: Leveraging Deep Learning and Explainable AI. *IEEE Access* **2024**, *12*, 153580–153595. [CrossRef]
23. Awad, A.I.; Babu, A.; Barka, E.; Shuaib, K. AI-Powered Biometrics for Internet of Things Security: A Review and Future Vision. *J. Inf. Security Appl.* **2024**, *82*, 103748. [CrossRef]
24. Sammoud, A.; Hamdi, O.; Chalouf, M.-A.; Montavont, N. The Contributions of Biometrics and Artificial Intelligence in Securing the IoT. In *Intelligent Security Management and Control in the IoT*; Wiley: Hoboken, NJ, USA, 2022; pp. 197–221. [CrossRef]
25. Alotaibi, A.; Aldawghan, H.; Aljughaiman, A. A Review of the Authentication Techniques for Internet of Things Devices in Smart Cities. *Sensors* **2025**, *25*, 1649. [CrossRef]
26. Cao, Q.; Shen, L.; Xie, W.; Parkhi, O.M.; Zisserman, A. VGGFace2: A Dataset for Recognising Faces Across Pose and Age. In *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition, Xi'an, China, 15–19 May 2018*; IEEE: Piscatway, NJ, USA, 2018; pp. 67–74. [CrossRef]
27. Chung, J.S.; Nagrani, A.; Zisserman, A. VoxCeleb2: Deep Speaker Recognition. In Proceedings of the Interspeech, Hyderabad, India, 2–6 September 2018; pp. 1086–1090. [CrossRef]
28. Maio, D.; Maltoni, D.; Cappelli, R.; Wayman, J.L.; Jain, A.K. FVC2004: Third Fingerprint Verification Competition. In Proceedings of the International Conference on Biometric Authentication, First International Conference, ICBA 2004, Hong Kong, China, 15–17 July 2004; pp. 1–7. [CrossRef]
29. Huang, G.B.; Mattar, M.; Berg, T.; Learned-Miller, E. *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments*; Technical Report 07-49; University of Massachusetts: Amherst, MA, USA, 2008.
30. Gemmeke, J.F.; Ellis, D.P.; Freedman, D.; Jansen, A.; Lawrence, W.; Moore, R.C.; Ritter, M. Audio Set: An Ontology and Human Labeled Dataset for Audio Events. In *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing. Where Signals Meet Intelligence, New Orleans, LA, USA, 5–9 March 2017*; IEEE: Piscatway, NJ, USA, 2017; pp. 776–780. [CrossRef]
31. Yang, R.; Zhang, Q.; Meng, L.; Wang, C.; Hu, Y. LUTBIO: A Comprehensive Multimodal Biometric Database Targeting Middle-Aged and Elderly Populations for Enhanced Identity Authentication. *Inf. Fusion* **2025**, *118*, 102945. [CrossRef]
32. Ross, A.; Jain, A.K. Multimodal biometrics: An overview. In Proceedings of the 12th European Signal Processing Conference (EUSIPCO), Vienna, Austria, 6–10 September 2004; pp. 1221–1224.
33. Jain, A.K.; Nandakumar, K.; Ross, A. Score normalization in multimodal biometric systems. *Pattern Recognit.* **2005**, *38*, 2270–2285. [CrossRef]
34. Bedari, A.; Wang, S.; Yang, W. A Secure Online Fingerprint Authentication System for Industrial IoT Devices Over 5G Networks. *Sensors* **2022**, *22*, 7609. [CrossRef]
35. Sarier, N.D. Privacy Preserving Biometric Authentication on the Blockchain for Smart Healthcare. *Pervasive Mobile Comput.* **2022**, *86*, 101683. [CrossRef]
36. Tran, Q.N.; Turnbull, B.P.; Wang, M.; Hu, J. A Privacy-Preserving Biometric Authentication System with Binary Classification in a Zero Knowledge Proof Protocol. *IEEE Open J. Comput. Soc.* **2022**, *3*, 1–10. [CrossRef]
37. Prasad, S.; Tiwari, N.; Chawla, M. Zero-Knowledge Proofs in Biometric Authentication Systems: A Review. In *Proceedings of the Congress on Smart Computing Technologies*; Springer: Singapore, 2024; pp. 281–295. [CrossRef]
38. Lai, J.; Wang, T.; Zhang, S.; Yang, Q.; Liew, S.C. BioZero: An Efficient and Privacy-Preserving Decentralized Biometric Authentication Protocol on Open Blockchain. *arXiv* **2024**, arXiv:2409.17509.
39. Yirga, T.G.; Gizachew Yirga, H.; Addisu, E.G. Cryptographic Key Generation Using Deep Learning with Biometric Face and Finger Vein Data. *Front. Artif. Intell.* **2025**, *8*, 1545946. [CrossRef] [PubMed]
40. Ouyang, R.; Wu, X.; Lv, Z. Personal Identification and Authentication in Multi-Task EEG Database Using EEGNet and Siamese Network. In Proceedings of the International Joint Conference on Neural Networks, Yokohama, Japan, 30 June–5 July 2024; pp. 1–8. [CrossRef]
41. Sharma, A.K.; Bhattacharya, S.; Reza, M. Dual Channel Multi-Attention in ViT for Biometric Authentication Using Forehead Subcutaneous Vein Pattern and Periocular Pattern. *arXiv* **2024**, arXiv:2412.19160.
42. Zheng, X.X.; Taha, B.; Rahman, M.M.U.; Masood, M.; Hatzinakos, D.; Al-Naffouri, T. Multimodal Biometric Authentication Using Camera Based PPG and Fingerprint Fusion. *arXiv* **2024**, arXiv:2412.05660. [CrossRef]
43. Balasubramaniam, S.; Kadry, S.; Prasanth, A.; Dhanaraj, R.K. *AI Based Advancements in Biometrics and Its Applications*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2024. [CrossRef]

44. Basaligheh, P.; Kothawade, S. Biometric Authentication Systems: Advances in Security and Usability. *Int. J. Recent Adv. Eng. Technol.* **2025**, *13*, 20–25.
45. Kasri, W.Y.-A. From Vulnerability to Defense: The Role of Large Language Models in Enhancing Cybersecurity. *Computation* **2025**, *13*, 30. [CrossRef]
46. Moustafa, M.T. Generative Fuzzer-Driven Vulnerability Detection in the Internet of Things Networks. *Appl. Soft Comput.* **2025**, *128*, 112973. [CrossRef]
47. Choi, H.; Jeong, J. Domain-Specific Manufacturing Analytics Framework: An Integrated Architecture with Retrieval-Augmented Generation and Ollama-Based Models. *Processes* **2025**, *13*, 670. [CrossRef]
48. Lu, Y.; Tong, L.I.U.; Xiang, Z.H.U. Research on Constructing a Network Security Event Knowledge Network Based on Multi-Source Data. In *Proceedings of the IEEE International Conference on Electronics and Devices, Computational Science (ICEDCS)*; IEEE: Piscatway, NJ, USA, 2024; pp. 1–6. [CrossRef]
49. Sharma, N.; Sikarwar, C. Harnessing the Power of Behavioral Biometrics and AI in Cyber Security. *Int. Res. J. Mod. Eng. Technol. Sci.* **2025**, *7*, 234–245.
50. Nasr, M.; Brzostowski, K.; Piórkowski, A.; Abd El-Samie, F.E. Cancelable Biometric Authentication Leveraging Empirical Mode Decomposition and Quaternion Representations for IoT Security. *Sci. Rep.* **2025**, *15*, 10997. [CrossRef]
51. Górski, T. Software Architecture Description in Original Software Publications. *Softw. Impacts* **2026**, *27*, 100802. [CrossRef]