



Biometric Arms Race: Multimodal Fusion Systems

The battle for identity at digital platforms is moving from passwords and single trait checks to multilayered security systems.

By **Ghulam Mustafa** | University of Agriculture Faisalabad | Published: August 2026

✉ mustafakhichi80@gmail.com | Position: Computer Science Student

Introduction

In today world, the most critical challenge faced by technology is making sure our devices and data remain secure, confidential and intact. The rise of technologies like Deep learning, Deepfake-audios, videos and images created by AI which mimics a real person appearance and voice and attacks such as spoof attacks-deceives users and gain access to sensitive data, these technologies has made it very easy for hackers to mimic fingerprint, facial pattern and overtook the system. This evolution results in failure and insecurity of traditional unimodal systems-require only one trait to gain access.

"Multimodal Fusion Systems combine two or more traits and creates a multilayer security system that forces an intruder to run different and unique locks simultaneously, resulting in nearly impossible spoofing."

Table 1: Comparative Analysis of Biometric Fusion Modalities credit: Jajan, K.I., & Abdulllah, W. M. (2026).

Biometric Traits	Fusion Level	Accuracy/EER	System Strength
Iris and Face	Score Level	100%	Highest accuracy with moderate overhead
Face and Fingerprint	Score Level	99.65%	Reliable performance across large populations
Iris and Fingerprint	Feature Level	99.4%	High data richness but complex to process
Gait and Face	Feature Level	98.75%	Strong defence against visual spoofing
Fingerprint and ECG	Feature Level	High (0.99 AUC)	Combines physiological and behavioural traits
Face and Speech	Decision Level	97.51%	Robust in uncertain or noisy environments

Technology Overview

To make our systems secure, the Tech industry has moved from unimodal to multimodal Fusion Systems that combines more than one biological trait-such as face recognition is paired with iris, veins pattern or walking style-to make security far more secure and robust. The “Degree of Freedom” has increased in the authentication layer and significantly raising the computational cost and complexity for any intruder.

How Fusion works? Fusion strategies are divided into two main categories: Fusion Before Matching and Fusion After Matching. **Fusion Before Matching:** It includes Sensor Level Fusion-collect data directly from hardware sources like sensors, before any features are extracted. This level is rich in data but due to high latency it has not been used in real time mobile apps. Feature Level Fusion-extracts specific mathematical vectors from each biological trait and links them into one high dimensional digital signature. **Fusion After Matching:** It includes



Match Score-Level Fusion-combines independent scores from all the sensors and assign weight to each score. It has high flexibility and accuracy of almost 100% in recent studies. Decision Level Fusion is a fast method where each biological trait makes and independent “Accept” or “Reject” call before a final majority vote is taken.

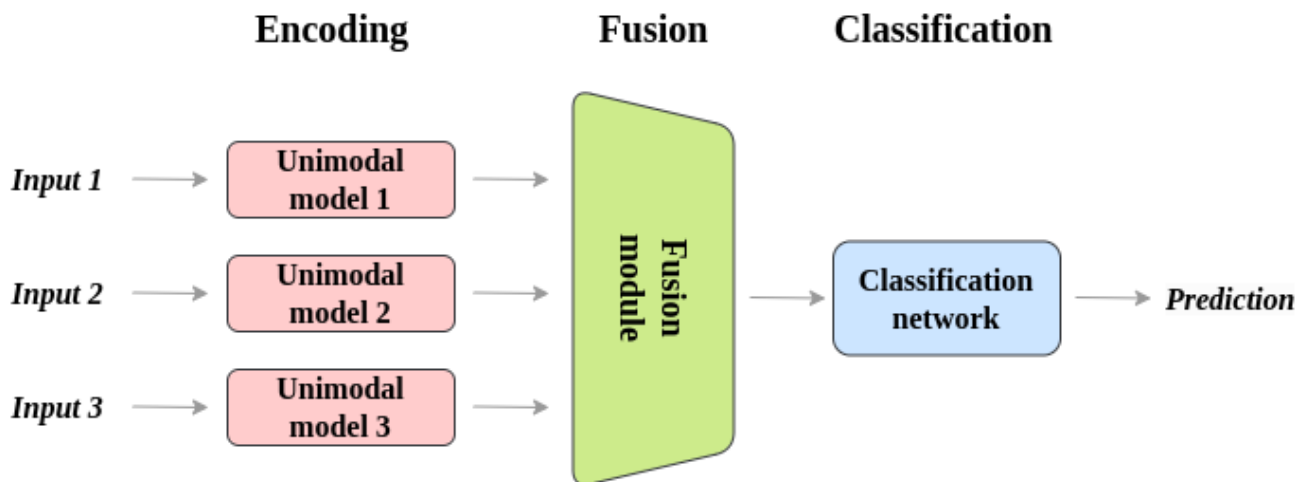


Figure 1: Multimodal System Architecture. Credit: image by roboflow

Key Findings & Impact

Finger print and Face recognition are no longer safe because rise in technologies like Deepfakes and spoofing which makes unimodal systems easy to hack. Multilayer defence is nearly impossible to spoof, it combines two or more traits (face + iris). It uses Algorithms like CNN (for images) and Transformers (for syncing your voice with your face) for security. Accuracy level of multimodal system is incredibly strong. E.g. iris and face scans can reach 100% accuracy. Multimodal systems have raised the cost of hacking and power of computing to break the system due to increase of “Degree of Freedom” in security. The End of “Black Box”: We are shifting towards Explainable AIs which tells us the reason if access is denied. Traditional systems are black box does not tell why the access is denied.

The Role of AI and the Future of Intelligent Security: In the current world of 2026, the effectiveness of a biometric system depends on the Artificial Intelligence that powers its backend. A system must be able to process a variety of inputs-like high-definition facial images and complex voice frequencies simultaneously without lagging to handle big data effectively. This is where advanced neural networks and optimization algorithms become essential. Modern systems have shifted from manual programming towards Deep Learning architectures that learn to identify users on their own. Convolutional Neural Networks are the primary tools for extracting features from image based traits like iris or fingerprint. VGG-16 are used to extract deep spatial features that remain perfect even if lighting is bad or user is moving. Temporal models are designed to remember patterns (rhythm of walk or signature of a person) over a time period. Transformer networks are good at “cross-modal relationships”, meaning they can sync facial expressions of users with their voice pitch to catch deepfakes. There has been a 40% increase in hybrid models since 2022. These systems don’t just rely on one type of fusion; for high security tasks they might use “Feature-level” fusion and “Score-level” fusion for faster and everyday logins. Adaptive fusion means if a user is in a dark room, the AI automatically ignores the camera data and puts 100% of its trust in the voice or fingerprint sensors instead. The most difficult part of the biometric arms race is about trust and authenticity. We are creating a super



identity for every user that creates a massive responsibility: if a password is stolen, it can be reset, but if a fused biometric profile of user is leaked, the user's biological identity is compromised for his entire life.

The Ethics of Algorithmic Bias: Many biometric AI models were trained in controlled labs environment with limited diversity, meaning they might not perform well when dealing with real world objects and for certain ages, or genders. A good system must be tested against real-world and datasets should be diverse to ensure it is not bias to anyone. Explainable AIs are vital for this purpose because they explain the reason behind every action they perform.

What's Next?

The winner of this Biometric arms race will not be the company with the most sensors, but the one with the most trusted and secure architecture. As we are moving towards "Context-Aware Authentication", where systems automatically adjust weights of traits according to the environment. The Future involves making these systems fair for everyone with the help of Explainable AI. Ultimately, Multimodal Fusion System ensures that identity of a user is no longer just a password, but a secure, authentic, real pattern of who we are.

About the Author

Ghulam Mustafa is a Computer Science student at the University of Agriculture Faisalabad (UAF). He is currently specializing in backend development. He can be reached at mustafakhichi80@gmail.com.

Keywords: Artificial Intelligence; Biometrics; Explainable AI; Black Box; Multimodal Fusion

References

- [1] Kaur, G., Bhushan, S., & Singh, D. (2017). Fusion in Multimodal Biometric System: A Review. Tech Journal.
- [2] Jajan, K. I., & Abdullah, W. M. (2026). Multimodal Biometric Fusion Strategies: A Comparative Review.).