Multimodal biometric authentication of identical twin users in e-Learning platforms: Security architecture

Sorin SOVIANY, Cristina-Gabriela GHEORGHE, Maria GHEORGHE-MOISII

National Institute for Research & Development in Informatics – ICI Bucharest,
Bucharest, Romania

sorin.soviany@ici.ro

Abstract: *This article explores the challenges of the identical twins' authentication in online learning environments, where the identity substitution can compromise the fairness of assessments. The paper proposes a multimodal and continuous authentication system that combines face (static biometric) together with voice and keystroke dynamics (behavioural biometrics), adapted to the subtle differences between twins. The research highlights the effectiveness of the proposed system in preventing frauds and aims to the design and development of efficient security solutions for e-Learning platforms, considering the biometric similarities of twin users.*

Keywords: Continuous authentication, Multimodal biometrics, Identical twins, e-Learning security, Feature fusion.

1. Introduction

The e-Learning platforms have become essential for modern education, ensuring flexibility and extended access to educational resources (courses, online assessments, various digital resources). They provide reliable support for scalable learning environments. As the number and relevance of online examinations grow, the integrity of user authentication becomes a critical requirement. The online assessments represent contexts in which the user identity must be reliably validated.

The identical twins represent a challenging authentication scenario as their biometric traits (faces, voices, certain behavioral patterns) could be extremely similar. Their high biometric similarity can effectively reduce the inter-class separability, increasing the false acceptance rates. These cases can defeat the traditional (password-based and single-biometric) authentication systems efficiency (Phillips et al., 2011), (Afaneh et al., 2017). In high-stakes e-Learning scenarios, one twin could attempt to take an exam on behalf of the other. For the identical twins, their nearly indistinguishable physical and sometimes behavioural characteristics in the use-cases of e-Learning platforms may significantly compromise the integrity and reliability of the online assessments, enabling the identity substitution and exam frauds.

https://doi.org/10.58503/icvl-v21y202638

The traditional authentication methods with passwords, PIN (Personal Identification Number) codes or tokens are vulnerable to sharing, data leakage, social engineering attacks. Another issue is the identity substitution between identical twins; the traditional methods fail to distinguish between identical twins, resulting in increased vulnerability to identity substitution. Biometric authentication, including modalities such as fingerprint, iris, face recognition, speaker verification, or keystroke dynamics, provides a more reliable link between the digital identity and the real person. However, the conventional face or voice recognition systems often fail to distinguish the subtle differences between twins (Sun et al., 2010), (Hollingsworth et al., 2011).

This paper proposes a twin-aware multimodal and continuous authentication architecture for e-Learning environments. The designed multimodal biometric system is able to distinguish identical twin users through a combination between face and behavioural biometrics (voice and keystroke). The design extends the previous multimodal biometric security model for e-Learning (Soviany et al., 2025), in which a biometric system architecture was proposed, with pre-classification fusion. The extension provides a dedicated feature-space design for twins, a Twin Similarity Factor (TSF) for adaptive decision thresholds, and a continuous Replacement/Risk Suspicion Index (RSI) for mid-session substitution detection.

The design looks to ensure continuous authentication for the users of e-Learning platforms, while combining the following modalities: *face*, *voice recognition* and *keystroke dynamic*, another behavioural biometric that can be completed with additional behavioural features (Hussain, 2025). The system not only authenticates the user at the beginning of the session, but also continuously throughout it, reducing the risk of identity substitution. The remainder of the paper has the following structure: Section 2 - Related works; Section 3 - System architecture and functional components; Section 4 - Experimental setup and performance; Section 5 - Conclusion and future works.

2. Related works

(Vevera et al., 2024) emphasize the importance of cybersecurity education and the integration of cyber diplomacy concepts into university curricula, highlighting the role of user awareness in preventing digital risks. This approach is relevant for the authentication of twin users on e-Learning platforms. The effective cybersecurity in digital education requires not only technical safeguards and trained professionals but also coherent and well-designed security frameworks able to protect the identity integrity in online learning and assessment environments. The risk of identity substitution and impersonation during online examinations highlights the need for advanced authentication solutions. This reasons the recent research on multimodal biometric systems, particularly for challenging scenarios such as the identical twin users in e-Learning. (Zamfir et al., 2023) highlight that the use of virtual educational environments involves the extensive collection of sensitive data, which imposes strict requirements regarding security, digital identity, and user control over personal information - issues that are directly relevant to the authentication of twin users on e-Learning platforms. (Badawi & Ciupercă, 2023) emphasize the importance of understanding both the human and technological dimensions in resilient educational systems in the context of global crises, arguing that online education must be integrated as a comprehensive service system.

The design of unimodal biometric authentication systems deals with significant challenges (Bowyer and Flynn, 2016). This involves the use of multimodal feature-fusion techniques to capture subtle discriminative patterns that are not visible in a single modality. Another work proposes an innovative method for the recognition of monozygotic twins by analysing the curvature area of the facial contour using Simpson's rule and ML (Machine Learning) algorithms, achieving superior performances for identification of subtle differences in 2D and 3D images (Sanil et al., 2025). (Kamlaskar & Abhyankar, 2021) proposed a multimodal iris–fingerprint biometric system based on an optimal feature-level fusion model, using Canonical Correlation Analysis to efficiently combine complementary information. (Mohammed & Hashim, 2020) proposed a multimodal biometric system designed to distinguish identical twins, using the Aspect United Moment Invariant descriptor to extract robust features and highlight individual uniqueness. Nsaif and Hasan propose an advanced biometric authentication method that combines detailed facial recognition, geometric and textural analyses, and dynamic facial micro-expressions; this allows to overcome the challenge of differentiating between identical twins (Nsaif & Hasan, 2025).

A systematic review of multimodal biometric methods used in the differentiation of monozygotic twins, conducted by (Kumari et al., 2025), highlighted the superiority of fingerprints and the complementary value of other biometric traits. (Sami et al., 2022) explain the difficulty of discriminating against identical twins using standard face recognition, despite of some recent technical advances based on deep CNN (Convolutional Neural Networks) together with twin-specific face benchmarks. These advances reinforce the conclusion that no single biometric modality is sufficient, reasoning the integration of complementary modalities. The multimodal biometric systems combining face, voice and behavioural traits have been shown to reduce false acceptance rates and increase robustness against spoofing, especially when feature-level or score-level fusion strategies are applied. (Herbadji et al., 2020) indicated that score-level or feature-level fusion significantly improves the recognition accuracy, as each modality captures independent traits; the combination of fingerprint and palmprint traits through feature-level fusion improves overall recognition accuracy in contactless settings. (Drozdowski et al., 2021) analysed the data fusion in facial recognition systems under privacy constraints, revealing that this process can be optimized for high accuracy even with partially anonymized data - an important issue in the GDPR context for e-Learning; they proposed advanced fusion strategies enabling efficient and privacy-preserving biometric indexing.

(Parde et al., 2022) show that the face and voice integration significantly increase the accuracy for the identical twin's identification; modern CNN architectures can surpass human performance when distinguishing twins under varying viewpoints, emphasizing the discriminant power of deep learning for twin identification; this confirms that twins can be distinguished using deep models if the images are sufficiently diverse and processed within discriminative subspaces. Multimodal twin-specific datasets (e.g. Audio-Visual Twins Database, CASIA twins' extensions) were considered for processing in (Akin, Kacar & Kirci, 2018) to evaluate the reliability of the multi-biometric approach even for the twin recognition case. Some subtle but stable differences present in the twins' speech signals could be exploitable for multimodal systems. It seems that voice contains unique features for each twin, reasoning for the integration of voice features into multimodal systems.

Hussain (2025) shows that the keystroke dynamics can detect identity substitution even when physical biometrics are similar; this work finds that the keystroke dynamics and interaction patterns can be effectively used for continuous authentication in e-Learning systems, helping to detect impersonation during assessments.

In (Lu et al., 2020) a continuous authentication framework based on free-text keystroke dynamics was proposed. The method combines CNN and RNN (Recurrent Neural Network) architectures to learn sequential typing patterns without requiring fixed input text; this enables unobtrusive and persistent user verification without disrupting learning activities in e-Learning use-cases. As relying on behavioural patterns rather than physical traits, this method is relevant for e-Learning environments ensuring a strong potential for distinguishing highly similar users, such as identical twins, where traditional biometric modalities often fail. The authors stated that the experimental results on large public datasets demonstrate low error rates -EER (Equal Error Rate) down to 2.36%. Platforms such as edBB-Demo (Daza et al., 2024) combine behavioural biometrics with webcam and audio analysis for continuous student monitoring, confirming that the static authentication is not sufficient for online exams; the edBB-Demo platform integrates facial, audio, and behavioural monitoring to support secure online examinations. These studies show that continuous monitoring can address the identity substitution risks that static biometric checks fail to detect.

As concerning *the biometric authentication for e-Learning applications*, under the extending adoption of online and virtual learning platforms, the security of the user identity has become a critical challenge. Several studies report vulnerabilities in traditional authentication methods (passwords or single biometric) when used in remote education scenarios (Huan et al., 2020). (Curran & Curran, 2021) reviewed biometric authentication techniques applicable to online learning, also defining multimodal solutions (such as face and iris) capable to mitigate identity frauds. Afolabi and Adagunodo (2018) proposed a crypto-biometric architecture integrating facial recognition and encryption for secure multimedia e-Learning environments. Ivanova et al. (2019) highlighted the importance of trust and reliability through the TeSLA system, which incorporates biometric and behavioural verification mechanisms.

The present work differs from the previous ones by explicitly addressing the authentication of identical twin users in e-Learning platforms, with a methodology combining multimodal biometrics (face and behavioural traits) and continuous authentication over the full duration of the learning/examination session.

3. System architecture and functional components

3.1 Problem statement and system requirements

Let's consider an e-Learning platform where students access courses and take online exams. A user account and its associated identity are intended to correspond to a unique person. In the case of identical twins, one twin may attempt to take an exam on behalf of the other. The problem statement and system requirements include:

- *threat model for identical twins in e-Learning*. The main threat is the identity substitution between twins. The components of the threat model are:
  - ➢ *Scenario*: twin A enrolled, twin B takes exam. Twin A is the legitimate enrolled student while twin B attempts to access or complete an assessment using twin A's credentials and similar biometric patterns;
  - ➢ *Attacker capabilities*: access to credentials, similar face/voice, potentially similar typing style. The main assumptions are: twin B has full knowledge of the account credentials, can be present on camera and microphone, exhibits facial and vocal traits very similar to those of the victim (twin A), may share similar behavioural patterns (typing speed, general navigation);
  - ➢ *Security goals:* G1- the system should accurately distinguish between twin A and twin B at enrolment and login with high confidence; G2- the system should detect mid-session substitution (continuous authentication); G3- the system should comply with privacy regulations (GDPR) for biometric data, providing such protection without excessive burden or intrusion on legitimate users.
- *functional and non-functional requirements*. The functional requirements: are F1- initial strong verification (initial multimodal authentication at login or exam start, combining face and voice); F2- continuous verification during the session- primarily based on behavioural biometrics and periodic checks of face and voice; F3- twin-aware decision logic that adapts thresholds when a potential twin is detected or known; F4-risk-based response for elevated suspicion (warning, request re-authentication, session lock). The non-functional requirements are: N1- latency constraints, N2- low intrusiveness, N3- scalability to large numbers of users and concurrent sessions, N4- privacy

by design and compliance, especially with respect to GDPR and biometric data protection;

- *performance metrics*, that include: standard biometric/ML metrics -TPR (True Positive Rate), FPR (False Positive Rate), ROC (Receiver Operating Characteristic), AUC (Area Under the Curve), EER; twin-specific metrics - TFAR (Twin False Accept Rate -probability to accept the wrong twin as the genuine user), TFRR (Twin False Reject Rate - probability to reject the correct twin); continuous authentication metrics (average time to detect a substitution, false alarm rate per session).

3.2 System architecture: the functional components and formal specifications

The basic security architecture is depicted in Figure 1. The major functional components are: *face recognition* (data acquisition, pre-processing and local fusion -intra-modal), *voice recognition* (acquisition, pre-processing and local fusion), *other behavioural biometrics* like keystroke (acquisition, pre-processing and local fusion), *global feature-level fusion* (inter-modal), *continuous authentication engine* and *classification module for the twin's identification*.

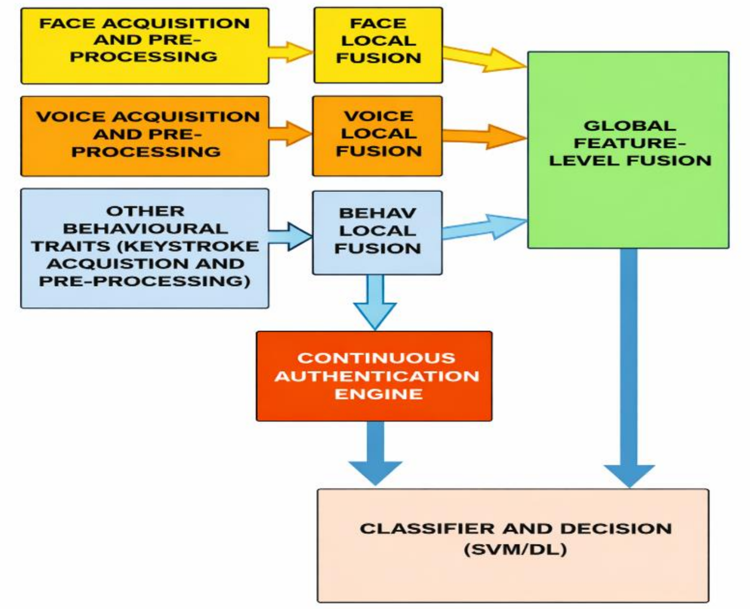
FACE ACQUISITION AND PRE-PROCESSING
FACE LOCAL FUSION
VOICE ACQUISITION AND PRE-PROCESSING
VOICE LOCAL FUSION
OTHER BEHAVIOURAL TRAITS (KEYSTROKE ACQUISTION AND PRE-PROCESSING)
BEHAV LOCAL FUSION
GLOBAL FEATURE-LEVEL FUSION
CONTINUOUS AUTHENTICATION ENGINE
CLASSIFIER AND DECISION (SVM/DL)

Figure 1. The security architecture for e-Learning applications

3.2.1 Formal problem specification

The problem specification considers 2 individuals having quite overlapped biometric data distributions X_b (for the biometrics b: F-face, V-voice, B-KS-behavioural keystroke and optionally other dynamic behavioural patterns): $X_F^{(i)}, X_V^{(i)}, X_{B-KS}^{(i)}, \mathrm{i} \in \{1,2\}$, where for the 2 twins the distances *D* between the face biometric data distributions satisfy the inequality:

$$D(X_F^{(1)}, X_F^{(2)}) << D(X_F^{(1)}, X_F^{(k)}), \forall k \neq 2 \tag{1}$$

3.2.2 The multimodal matching score with adaptive weighting

For this design one can consider the following definition of the multimodal matching score:

$$S = w_F \cdot S_F + w_V \cdot S_V + w_{B-KS} \cdot S_{B-KS} \tag{2}$$

The weights wb of the individual scores per biometric (Sb, $b \in \{F, V, B-KS\}$) are adjusted according to

$$w_b = \frac{1/\sigma_b}{\sum_{k \in \{F,V,B-KS\}} (1/\sigma_k)} \tag{3}$$

σ measures the score variability for a certain biometric modality between the 2 twins. The design assumes that a decrease in the differences between twins' per-modality scores corresponds to an increased contribution of behavioural components.

3.2.3 Processing components

The full processing pipeline contains *client-side* and *server-side components*. The *client-side components* include:

- the data acquisition and pre-processing for each modality: face (face detection, alignment, embedding), voice (MFCC extraction- Mel-Frequency Cepstral Coefficients providing voice features), other behavioural traits (keystroke events- key down/up timestamps, mouse trajectories). There will be: camera for face capture, microphone for voice capture, interaction monitoring for keystrokes and mouse movements, local pre-processing (face alignment, audio framing, keystroke event logging);
- the local anonymization/pseudonymisation where possible (for GDPR compliance)- secure channel to send features or pseudonymised features to the server.

The server-side components include:

- the feature normalization (e.g. with min-max or sigmoid functions) and storage (for the biometric templates);
- the per-modality scoring classifiers or matchers (similarity/ classifier output);
- the feature- or score-level fusion engine, with adaptive weighting for twins;

- the continuous monitoring/risk scoring and Replacement Suspicion Index module-RSI (sliding time windows);
- the access control and integration with e-Learning platforms.

3.3 Data modelling and feature space design

3.3.1 Feature space per modality

The feature vectors per biometric modality are specified as follows:

1) for the face feature space: $x_F \in \mathbb{R}^{d_F}$ the face feature vector with the dimensionality dF, with CNN embedding or textural/Haralick descriptors as in the previous work (Soviany et al., 2025). The Haralick features are based on co-occurrence matrices (Theodoridis & Koutroumbas, 2009);

2) for the voice feature space: $x_V \in \mathbb{R}^{d_V}$ the voice feature vector with the dimensionality dV, with MFCC features;

3) for the other behavioural feature spaces: $x_{B-KS} \in \mathbb{R}^{d_{B-KS}}$ the feature vectors for the other behavioural traits, with the dimensionality dB-KS (keystroke timing statistics and RNN-derived).

The design includes the feature space transformations for dimensionality reduction with PCA (Principal Component Analysis) /normalization. This process leads to the following feature vectors:

$$\hat{x}_F = PCA_F\left(x_F\right),\ \hat{x}_V = PCA_V\left(x_V\right),\ \hat{x}_{B-KS} = PCA_{B-KS}\left(x_{B-KS}\right) \tag{4}$$

The reduced dimensionalities are d_F^0, d_V^0 and d_{B-KS}^0, respectively. Each enrolled user U is associated with enrolment samples for each modality. A template is computed as the mean of the reduced feature vectors over enrolment samples. Therefore, the biometric (face, voice, other behavioural) templates for the user U are denoted as $\hat{m}_b^{(U)}, \mathrm{b} \in \{F, V, B-KS\}$.

3.3.2 Twin Similarity Factor (TSF)

Given a certain user U1 of the e-Learning platform, together with the potential twin U2, the design defines the amount TSF (Twin Similarity Factor):

$$TSF^{(U)} = \alpha \cdot D_F + \beta \cdot D_V + \gamma \cdot D_{B-KS} \tag{5}$$

D_b are the distances between twins for the given trait. The inter-twin distances per modality are given by

$$D_F = d\left(\hat{x}_F^{(U1)}, \hat{x}_F^{(U2)}\right) \tag{6}$$

$$D_V = d\left(\hat{x}_V^{(U1)}, \hat{x}_V^{(U2)}\right) \tag{7}$$

$$D_{B-KS} = d\left(\hat{x}_{B-KS}^{(U1)}, \hat{x}_{B-KS}^{(U2)}\right) \tag{8}$$

For the biometric template computed as the mean of the reduced feature vectors over enrolment samples, the inter-twin distance is

$$D_b = d\left(\hat{m}_b^{(U1)}, \hat{m}_b^{(U2)}\right), \mathrm{b} \in \{F, V, B-KS\} \tag{9}$$

These distances are weighted per modality (α, β, γ) based on their specific performances. For the distance measure one can consider the cosine or Euclidian distance. The weigths are chosen based on cross-validation per modality. For high similar twins one can expect to have small inter-twin distances per modality and therefore small TSF. In general, identical twins exhibit smaller inter-template distances than unrelated users. A small TSF indicates very similar twins and justifies constrained decision thresholds. In scenarios where twin information is not explicitly known as a priori, TSF can be estimated by comparing the enrolled user against a set of candidate impostors and identifying highly similar pairs.

3.3.3 Adaptive decision thresholds

As concerning the decision thresholding, the design specifies the adjustment of the baseline discrimination threshold θ_0 for a user U having a twin. TSF should be used to properly tune the discrimination thresholds θ:

$$\theta^{(U)} = \theta_0 - \lambda \cdot TSF^{(\mathrm{U})} \tag{10}$$

where $\lambda > 0$ is a scaling factor. The scaling coefficient λ is applied based on specific performances per modality to aggregate the best contributions for accurate discrimination.

This represents the thresholding adaptation for a user U with a known or detected twin. The baseline threshold should be optimized for overall verification performance on non-twin users (e.g., minimizing EER). The significance is that for more similar twins one should specify a more constrained thresholding, given the requirement to have a larger and clearer separation. Lower TSF (more similar twins) leads to more constrained thresholds, reducing the probability of false acceptance of the twin at the cost of potentially higher false rejections.

3.3.4 Data fusion (post-/pre-classification)

The data fusion process can be applied either post- or pre-classification, with score-level or feature-level rule, respectively. The feature-level fusion requires to ensure a certain homogeneity of the feature vectors allowing to apply a functional fusion method avoiding the concatenation. Homogeneity will require at least a common dimensionality of the original feature spaces. For image-based feature extractors (face recognition), a common feature extractor with textural Haralick features could provide homogeneity. The data fusion rule uses the weighted sum rule for both cases, score or feature fusion.

For the post-classification (score-level) fusion, given a certain biometric sample for the discrimination and having per-modality normalized scores S_F, S_V, S_{B-KS} ($S_F, S_V.S_{B-KS} \in [0,1]$), the fused score is computed using equation (2). The per-modality scores can be interpreted as the probability that the sample is genuine. The weighting could be either static -with weights empirically fixed per dataset, or data-driven – inverse-variance weightting according to equation (3), in which $\sigma_b, b \in \{F, V, B-KS\}$ is the score variance between the 2 twins per modality.

A practical approach is to derive weights from the variance of scores on twin impostor trials. Let the variances of impostor scores when the impostor is the twin: $\sigma_b^2, \mathrm{b} \in \{F, V, B-KS\}$. A lower variance (better separation between genuine and twin impostors) indicates stronger discriminative power. One may define the weight:

$$w_b = \frac{1/\sigma_b^2}{\sum_{k \in \{F,V,B-KS\}} 1/\sigma_k^2} \tag{11}$$

The modalities more robust to twins will receive higher fusion weights.

For the pre-classification (feature-level) fusion, the fusion rule is quite similar, therefore a weighted sum, but this should apply on feature vectors under their homogeneity conditions being assumed. Giving the transformed feature vectors - equations (4), with common dimensionality, the fusion rule provides the common feature vector *x*:

$$x = \sum_{b \in \{F,V,B-KS\}} W_b \cdot \hat{x}_b \tag{12}$$

3.3.5 Continuous risk scoring

At time step t, one can consider the windowed fused score S^t - a sequence of fused scores $S^t \in [0,1], t = 1, 2, \ldots T$ (or fused feature vector x^t). The fused scores are obtained at periodic intervals or upon specific events. The residual vs. enrolled template score is defined as

$$r^t = 1 - S^t \tag{13}$$

which represent deviation from the ideal genuine behaviour. A quite similar measure can be defined for the fused feature vectors, but only assuming normalized features (with sigmoid functions).

The Replacement Suspicion Index (RSI) over the last N windows (slots), at time t, is computed over a sliding window of length N:

$$RSI^t = \frac{1}{N} \cdot \sum_{i=t-N+1}^{t} r^i \tag{14}$$

This amount is used to estimate the risk of substitution between the twins. The suspicion of substitution is raised if the following condition is met: $RSI^t > \rho$, where ρ is a threshold value. If RSI^t exceeds a threshold ρ, the system flags a possible substitution event. This can be interpreted as continuous statistical monitoring of cumulative evidence that the current user behaves differently from the enrolled profile.

4. Experimental setup and performance. Discussion

The experimental protocol, either for synthetic data but also for real data, includes the following operations:

1. Split per-user: enrollment vs test (e.g. 2 images/audio samples for enrollment, rest for test);
2. Verification experiments: genuine trials (user vs self), impostor trials (non-twin impostors, twin impostors- twin A masquerading as twin B);
3. Evaluation: per modality (face-only, voice-only, behaviour-only), fused (face+voice, face+behaviour, full face+voice+behaviour).

The metrics to be used include ROC curves for each modality and for the fused case, accuracy, EER for each biometric and fused case, TFAR/TFRR, average time to detection in continuous scenario (for behaviour and periodic sampling for face and voice).

Since comprehensive public datasets containing multimodal biometrics for identical twins in e-Learning contexts are not readily available, the initial experiments will use synthetic data to emulate key properties of twin and non-twin users. The simulation will consider: N_p pairs of identical twins, N_s additional non-twin users, enrolment and raw samples per user for each modality. For each user and modality, a mean feature vector is drawn from a multivariate Gaussian. For twins, the 2 users share a similar mean (small offset), while non-twins have more separated means. Enrolment and raw samples are generated by adding Gaussian noise around the respective user means.

SVM (Support Vector Machine) classifiers with RBF (Radial Basis Function) kernels are trained on the verification pair datasets for each modality separately. Per-modality match scores are taken as the predicted probabilities for the genuine class. The twin-specific error rates are estimated by focusing on impostor pairs where the impostor corresponds to the twin of the genuine user, and computing TFAR at a default threshold (e.g., score ≥ 0.5). The synthetic results show lower EER for face and voice compared to other behavioural biometrics, non-zero TFAR for all modalities- with the highest TFAR on face or voice and other behavioural biometric depending on simulation parameters, higher TFAR for twins than for generic impostors, reflecting the challenge of twin discrimination. Table 1 shows preliminary results on a synthetic dataset simulating distributions quite similar with those of the previously referred datasets. For the same data, the accuracy is depicted in Figure 2.

Table 1. Preliminary results on synthetic data

Configuration	*EER (%)* *non-twins*	*TFAR (%)* *twins*	*TFRR (%)* *twins*
Face only	1.8	8.5	4.2
Voice only	2.1	6.9	5.0
Behaviour only	3.5	10.0	6.0
Face + Voice	0.9	3.2	2.5
Face + Voice + Other behavioural traits	0.7	1.8	1.9

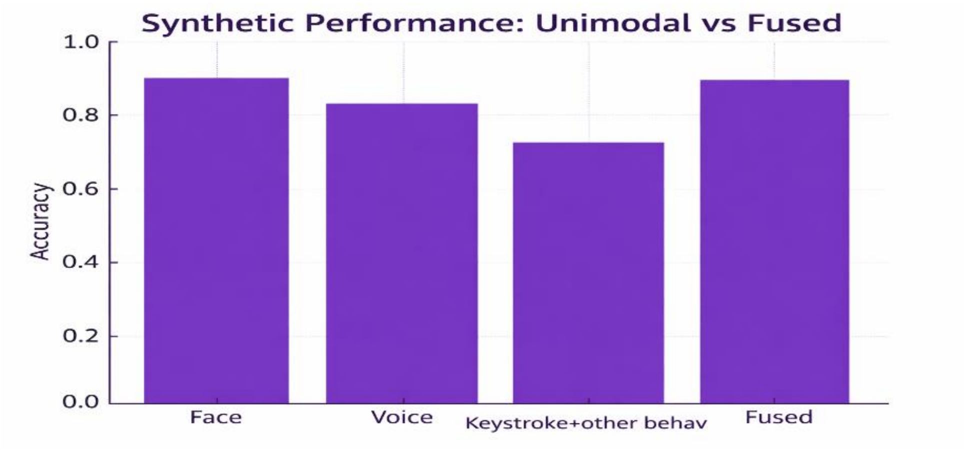

Figure 2. Accuracy for synthetic data

For the continuous authentication, the simulation should consider 2 time series of scores: a genuine session (where all scores are drawn from a high-mean, low-variance distribution) and a twin substitution session (where scores are genuine-like until a certain time point and then follow a lower-mean, higher-variance distribution-representing the twin impostor). RSI is computed over a sliding window, and its evolution must be compared in the 2 scenarios. For genuine sessions, RSI remains below a moderate threshold, whereas in the twin substitution scenario, RSI increases and crosses the threshold shortly after the substitution point. This demonstrates how RSI can be used to raise alerts and trigger additional checks during the session.

The main advantages of the proposed system (twin-aware multimodal architecture) are: improved security-robustness against twins, adaptive thresholds and weights, continuous authentication, GDPR compliance, scalability and adaptability. The limitations of the proposed system are: synthetic evaluation, dataset availability, implementation costs, dependency on the data quality, computational and deployment complexity, ethical concerns and user perception, user acceptance and privacy concerns.

5. Conclusions and future works

This paper addressed the challenge of authenticating identical twin users in e-Learning platforms. It proposes a twin-aware multimodal and continuous authentication architecture that combines facial, vocal, and behavioural biometrics, extending previous work on multimodal security models for e-Learning. The present work provides an integrated authentication methodology combining multimodal biometrics and continuous authentication for identical twins in e-Learning, showing its benefits but also some weak points for potential improvements. The *specific contributions* are: a twin-aware multimodal biometric system architecture, with specific design for identical twins in online learning environments; a basic mathematical foundation of the feature space- a feature space design, together with an adaptive fusion scheme and similarity between twin and non-twin users; an experimental protocol using synthetic data (simulation), with performance metrics estimation and illustrative results.

The steps toward full implementation include real-world testing (deploying the system in universities to validate its performance and collect empirical data), optimization of the multimodal algorithms (using Artificial Intelligence for continuous adaptation to individual behaviour and to reduce false positives), user education (to increase the trustness and acceptability), extending the application areas.

The more concrete future works include collecting a dedicated multimodal twin-e-Learning dataset (face, voice, keystroke), exploring more advanced fusion methods- such as neural networks with attention mechanisms or meta-learning, to learn fusion strategies directly from data, implementing formal privacy-preserving mechanisms, evaluating user perception and acceptance of continuous biometric authentication in educational environments.

Acknowledgments

The work presented in this paper is supported by the Core Program within the National Research Development and Innovation Plan 2022-2027, carried out with the support of MCID, project no. 23380601, "Advanced research in the Metaverse and emerging technologies for the digital transformation of society".

REFERENCES

Afaneh, A., Noroozi, F. & Toygar, Ö. (2017) Recognition of identical twins using fusion of various facial feature extractors. *EURASIP Journal on Image and Video Processing*. 2017, 81. https://doi.org/10.1186/s13640-017-0231-0.

Afolabi, A.O. & Adagunodo, E.R. (2018) Securing E-Learning System with Crypto-Biometric Multimedia. *Biostatistics and Biometrics Open Access Journal*. 7(2), 31-38. https://doi.org/10.19080/BBOAJ.2018.07.555710.

Akin, C., Kacar, U. & Kirci, M. (2018) A multi-biometrics for twins identification-based speech and ear. *arXiv preprint. arXiv:1801.09056*. https://arxiv.org/pdf/1801.09056 [Accessed: 17th November 2025].

Badawi, S., Ciupercă, E. M. (2023) Empowering education: learning from crises to achieve resilient education in Romania. *International Conference on Virtual Learning*, vol. 18, pp. 104-112, 2023. https://doi.org/10.58503/icvl-v18y202308.

Bowyer, K.bW. & Flynn, P.B.J. (2016) Biometric identification of identical twins: A survey. *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS), Niagara Falls, NY, USA, 2016*. pp. 1-8, https://doi.org/10.1109/BTAS.2016.7791176.

Curran, J. & Curran, K. (2021) Biometric Authentication Techniques in Online Learning Environments. In *Research Anthology on Developing Effective Online Learning Courses,* edited by Information Resources Management Association, IGI Global Scientific Publishing. 2021, pp. 867-879. https://doi.org/10.4018/978-1-7998-8047-9.ch042.

Daza, R., Morales, A., Tolosana, R., Gomez, L. F., Fierrez, J. & Ortega-Garcia, J. (2024) edBB-Demo: Biometrics and Behavior Analysis for Online Educational Platforms. *Proceedings of the AAAI Conference on Artificial Intelligence*. 37(13), 16422-16424. https://doi.org/10.1609/aaai.v37i13.27066.

Drozdowski, P., Stockhardt, F., Rathgeb, C., Osorio-Roig, D. & Busch, C. (2021) Feature Fusion Methods for Indexing and Retrieval of Biometric Data: Application to Face Recognition with Privacy Protection. *IEEE Access*, vol. 9, pp. 139361-139378, 2021. https://doi.org/10.1109/ACCESS.2021.3118830.

Hasan Nsaif, A. & Abduladheem Hasan, R. (2025) Next-Generation Biometric Authentication: Overcoming the Twin Identification Challenge with Advanced Facial Recognition and Multi-Modal Analysis Techniques. *Journal of Al-Qadisiyah for Computer Science and Mathematics,* 17(3), 133–152. https://doi.org/10.29304/jqcsm.2025.17.32383.

Herbadji, A., Guermat, N., Ziet, L., Akhtar, Z., Cheniti, M. & Herbadji, D. (2020) Contactless multi-biometric system using fingerprint and palmprint selfies. *Traitement du Signal*. 37(6), 889-897. https://doi.org/10.18280/ts.370602.

Hollingsworth, K. et al. (2011) *Genetically Identical Irises Have Texture Similarity That Is Not Detected by Iris Biometrics*. https://www3.nd.edu/~kwb/Hollingsworth_EtAl_CVIU_2011.pdf [Accessed: 17th November 2025].

Huan, L. Q., Nguyen, D. M., Pham, H. A. & Huynh-Tuong, N. (2020) Authentication in E-learning systems: Challenges and Solutions. *Journal of Engineering and Technology*. 3(SI1), SI95-SI101. https://doi.org/10.32508/stdjet.v3iSI1.516.

Hussain, S. (2025) Behavioral Biometrics and Continuous Authentication in Cybersecurity Systems https://www.researchgate.net/publication/392623351 Behavioral_Biometrics_ and_Continuous_Authentication_in_Cybersecurity_Systems doi: 10.13140/RG.2.2.35971 .41763 [Accessed: 8th December 2025].

Ivanova, M., Bhattacharjee, S., Marcel, S., Rozeva, A. & Durcheva, M. (2019) Enhancing trust in eassessment-the tesla system solution. *arXiv preprint arXiv:1905.04985*.

Kamlaskar, C. & Abhyankar, A. (2021) Iris-Fingerprint multimodal biometric system based on optimal feature level fusion model[J]. *AIMS Electronics and Electrical Engineering*. 5(4), 229-250. https://doi.org/10.3934/electreng.2021013.

Kumari, S., Bhatanagar, A., Agarwal, A., Kakkar, A. & Gupta, N. (2025) Decoding the Twin Code: Exploring Multimodal Biometrics in Identical Twin Differentiation – A Systematic Review. *Journal of Neonatal Surgery*. 14(8s), 980–989.

Lu, X., Zhang, S., Hui, P. & Lio, P. (2020) Continuous authentication by free-text keystroke based on CNN and RNN. *Computers and Security*. 96, Article 101861. https://doi.org/10.1016/j.cose.2020.101861.

Mohammed, B. O. & Hashim, S. Z. M. (2020) Individuality representation using multimodal biometrics with Aspect United Moment Invariant for identical twins. *Journal of Theoretical and Applied Information Technology*. 98(12), 2148–2157.

Parde, C. J. et al. (2022) Twin identification over viewpoint change: A deep convolutional neural network surpasses humans. *arXiv preprint*. https://arxiv.org/abs/2207.05316 [Accessed 17 Nov. 2025].

Phillips, P. J. et al. (2011) *Distinguishing Identical Twins by Face Recognition*. https://www3.nd.edu/~kwb/Phillips_EtAl_FG_2011.pdf [Accessed: 17th November 2025].

Sami, S. M., et al. (2022) Benchmarking human face similarity using identical twins. *IET Biometrics Journal*. 11(5), 459–484. https://doi.org/10.1049/bme2.12090.

Sanil, G., Prakasha, K., Prabhu, S. & Nayak, V. C. (2025) Area-based face curve characteristic analysis to recognize Multimodal 2D/3D monozygotic twins using Simpson's rule and Machine Learning. *Systems and Soft Computing*. 7, 200267. https://doi.org/10.1016/j.sasc.2025.200267.

Soviany, S., Gheorghe, C. & Gheorghe-Moisii, M. (2025) A security model with biometric components for e-learning systems. *Proceedings of the International Conference on Virtual Learning*, vol. 20, pp. 103-115. https://doi.org/10.58503/icvl-v20y202509.

Sun, Z., Paulino, A. A., Feng, J., Chai, Z., Tan, T. & Jain, A. K. (2010) A study of multibiometric traits of identical twins. *Biometric technology for human identification Vii*. Vol. 7667, pp. 283-294. SPIE https://doi.org/10.1117/12.851369.

Theodoridis, S. & Koutroumbas, K. (2009) Pattern Recognition 4th edition, Academic Press Elsevier.

Vevera, A.-V., Cîrnu, C.-E., Vasiloiu, I.-C. (2024) Enhancing cyber security education in Romania: integrating cyber diplomacy concepts into universities curricula. *Proceedings of the International Conference on Virtual Learning*, vol. 19, pp. 53-64. https://doi.org/10.58503/icvl-v19y202405.

Zamfir, M., Marinescu, I. A., Iordache, D., Barbu, M. & Cîrnu, C. E. (2023) Exploring ethical considerations in Metaverse from the education perspective. *International Conference on Virtual Learning*, vol. 18, pp. 91-100, 2023. https://doi.org/10.58503/icvl-v18y202307.