
Journal of Informatics and Web Engineering

Vol. 5 No. 3 (October 2026)

eISSN: 2821-370X

Mask-Resilient Multi-Biometric Authentication for Healthcare

Caleb Yeo¹, Siew-Chin Chong^{2*}, Kian-Ming Lim^{3}, Lee-Ying Chong⁴**

^{1,2,5}Faculty of Information Science and Technology, Multimedia University, Jalan Ayer Keroh Lama,
75450 Bukit Beruang, Melaka, Malaysia

³Faculty School of Computer Science, University of Nottingham Ningbo China, Zhejiang Province, 315100, China

*corresponding author: (chong.siew.chin@mmu.edu.my; ORCID: 0000-0003-0421-4367)

**corresponding author: (Kian-Ming.Lim@nottingham.edu.cn; ORCID: 0000-0003-1929-7978)

Abstract - The fast integration of digital healthcare platforms, including telemedicine and electronic health records, has created a significant need for secure, contactless authentication methods. Traditional login methods such as passwords and PINs are increasingly insufficient for ensuring the security of sensitive patient data. This paper presents a robust, lightweight Healthcare Authentication System that five biometric modules to address major challenges in facial recognition: spoofing attacks, liveness detection, and facial occlusion due to mask-wearing. The system leverages RetinaFace for face detection, a MobileNetV2 model for mask classification, a lightweight CNN for blink-based liveness verification, and MiniFASNet for spoof detection. For identity matching, the system utilizes InceptionResNetV1 with ArcFace loss, enabling effective recognition of both masked and unmasked faces through a dual-embedding strategy. Each module was evaluated using benchmark datasets: CASIA-FASD for spoofing, RMFD for mask detection, open/closed eye datasets for blink detection, and VGGFace2 (with synthetic masking) for recognition. The system achieved 91% accuracy in blink detection, 85% in spoof detection, 97% in mask classification, and 96% in face recognition. The complete pipeline achieved 88% overall accuracy on the evaluated dataset. The proposed solution offers a practical and scalable approach for secure biometric authentication in healthcare environments, particularly in clinical settings where hygiene, user accessibility, and anti-spoof measures are essential. This work provides a foundation for future research in offline biometric verification, depth-based spoof detection, and multilingual interface design, contributing toward more inclusive and secure healthcare technologies. While we do not fine-tune model backbones, our contribution lies in a clinically viable integration strategy that combines blink-based liveness, anti-spoofing, mask detection, and face verification under a fixed operating point with identity-disjoint evaluation, achieving 88% end-to-end accuracy on public-dataset inputs.

Keywords—Blink Detection, Anti-Spoofing, Masked Face Recognition, Biometric Security, Healthcare Authentication

Received: 21 November 2025; Accepted: 20 March 2026; Published: 16 October 2026

This is an open access article under the [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



1. INTRODUCTION

The fast-growing adoption of digital technologies in the healthcare sector has fundamentally transformed the way medical services are accessed, delivered, and managed. With the increasing reliance on platforms such as telemedicine and Electronic Health Records (EHR), there is a pressing need for robust and secure identity verification systems [1].



Journal of Informatics and Web Engineering

<https://doi.org/10.33093/jiwe.2026.5.3.2>

© Universiti Telekom Sdn Bhd.

Published by MMU Press. URL: <https://journals.mmupress.com/jiwe>

Ensuring that only authorized individuals can access sensitive medical data is critical not only for compliance with data protection regulations such as HIPAA but also for maintaining patient trust and safety.

Historically, healthcare systems have relied on traditional authentication mechanisms such as passwords, PINs, or smart cards. However, these methods are increasingly recognized as insufficient in countering modern cybersecurity threats. They are prone to being forgotten, stolen, guessed, or phished, which opens the door to identity theft and unauthorized access to confidential patient records [2]. Furthermore, in fast-paced clinical environments, the need for quick and contactless authentication solutions has become more apparent, especially in the wake of the COVID-19 pandemic.

Biometric authentication, particularly face recognition, offers a compelling solution due to its non-intrusive nature and high user acceptance. Nevertheless, conventional face recognition systems face significant limitations when deployed in real-world hospital settings. Challenges such as spoofing attacks using photos or videos and the presence of face masks—now commonly worn for infection control—can drastically reduce recognition accuracy and system reliability [3], [4], [5].

To address these challenges, this paper proposes an enhanced Healthcare Authentication System that integrates three key components: blink detection for liveness verification, face mask detection for robustness under occlusion, and anti-spoofing mechanisms to counter fraudulent access attempts. By combining these biometric strategies with advanced deep learning models, the system ensures a higher degree of reliability and security in patient identity verification. The system is designed for lightweight operation and is suitable for integration into existing hospital workflows. The remainder of this paper discusses related work, system architecture, evaluation methodology, results, and potential directions for future enhancement.

This work contributes a mask-resilient face authentication pipeline tailored to healthcare constraints: (i) a liveness-gated(blink) and anti-spoofed verification flow, (ii) a dual-embedding enrolment/recognition strategy to handle masked vs. unmasked faces, (iii) a transparent operating point ($\text{Blink} \geq 0.7$, $\text{anti-spoof} = \text{Real}$, $\text{CosSim} \geq 0.75$) with identity-disjoint evaluation on public datasets, and (iv) a lightweight, deployable implementation validated end-to-end on commodity hardware (MacBook Air M1). We find that this integration-first design achieves 88% pipeline accuracy while prioritising low FAR, aligning with practical clinical requirements.

2. LITERATURE REVIEW

2.1 Biometric Authentication in Healthcare

Biometric systems in healthcare improve both convenience and security by allowing seamless identity verification. Studies show the growing adoption of facial recognition over fingerprints in clinical settings due to its contactless nature [1]. However, challenges like facial occlusion and spoofing attempts require multi-layered verification methods [2].

2.2 Spoof Detection and Liveness Verification

Presentation attacks involve using photos, videos, or masks to trick facial recognition systems. Anti-spoofing techniques such as MiniFASNet use CNN architectures to detect texture differences between real and spoof faces [3]. Liveness detection using blink detection or head motion provides an additional security layer and improves reliability in practical scenarios [4].

2.3 Masked Face Recognition

The COVID-19 pandemic necessitated the development of models that can recognize faces even when partially covered by masks. Researchers proposed training on masked datasets and using embedding techniques like ArcFace to support both masked and unmasked comparisons [6]. Dual embeddings, as explored in recent dual-proxy network approaches, can effectively improve recognition accuracy under occlusion [7].

2.4 Comparison with Existing Systems

Many commercial offerings (e.g., FaceTec, Onfido) provide biometric verification, but they typically do not publish open academic datasets or healthcare-specific evaluations. Recent academic prototypes often target single challenges such as blink or mask detection [4], [8], [9], [10], while few integrate spoof, blink, mask awareness, and face verification into a single end-to-end pipeline [11].

Table 1. Comparison with Prior Multi-biometric / Mask-resilient Work

Work (Year)	Components (Blink / Spoof / Mask / FR)	Data and protocol	Operating point transparency	End-to- end pipeline	Reported end- to-end outcome
Atoum et al. (2017) [3]	X / ✓ / X / X	Public anti-spoof sets; module-level	—	X	—
“Blinking for Photo Attacks” (2019) [4]	✓ / X / X / X	Blink module dataset; module-level	—	X	—
FRVT 6A (2020, NIST) [5]	(evaluation programme)	Large-scale eval; vendor submissions	Yes (protocol is explicit)	X (not a single pipeline)	Benchmarks across vendors (not your setup)
Montero et al. (2022) — Multi-Task ArcFace (masked FR) [6]	X / X / (aux) / ✓	Masked-FR benchmarks; module-level	Sometimes τ /EER; pipeline OP not central	X	—
Shuhui and Xiaochen (2022) — Dual-Proxy MFR [7]	X / X / (task) / ✓	Masked-FR benchmarks; module-level	Sometimes τ /EER; pipeline OP not central	X	—
Himeur et al. (2023) — Mask detection in smart cities [8]	X / X / ✓ / X	Mask-det datasets; module-level / app survey	—	X	—
Qin et al. (2022) — Survey on identity fusion [11]	(review)	Survey	—	X	—
CDCN for FAS (2020) [12]	X / ✓ / X / X	Anti-spoof benchmarks; module-level	—	X	—
MobileNetV2 (2018) [13]	(backbone)	—	—	X	—
Inception-ResNet / ArcFace (2017–2019) [14][15]	(backbone/loss for FR)	FR benchmarks	Sometimes τ /EER per FR protocol	X	—
Masked Face datasets / studies (2021–2023) [16][17]	(data / task focus)	Dataset papers; evaluation protocols per set	—	X	—
Mobile DL accel (2016) [18] and ShuffleNet / mobile inference (2018) [20]	(efficiency)	—	—	X	—
RetinaFace (2020) [19]	(detector)	Detector benchmarks	—	X	—
Ours (this work)	✓ / ✓ / ✓ / ✓	Public datasets; identity-disjoint hold-out (notebook-execution)	Yes (Blink $\geq$ 0.7; anti-spoof = Real; CosSim $\geq$ 0.75)	✓	88% overall

Note: “—” denotes not reported or not applicable (module-level). “End-to-end outcome” refers to a full-pipeline accept/reject decision rather than single-module metrics.

Table 1 summarizes how representative prior efforts relate to our scope. Most cited works address single modules (e.g., blink [4], anti-spoof [3], [12], masked FR [6], [7], [16], [17]) or provide datasets/surveys [11], [5], rather than an integrated pipeline that jointly enforces blink + anti-spoof + mask-aware verification. In contrast, our contribution is an integration-first pipeline with a transparent operating point (Blink ≥ 0.7 ; anti-spoof = Real; CosSim ≥ 0.75) and an identity-disjoint hold-out evaluation on public datasets, yielding 88% end-to-end accuracy.

3. RESEARCH METHODOLOGY

3.1 System Design

The proposed Healthcare Authentication System integrates five biometric modules to ensure secure, lightweight identity verification suitable for hospital environments. The system incorporates:

- **Face Detection:** RetinaFace is employed to locate facial regions with high precision by leveraging a single-shot detection architecture and facial landmark localization [19]. It enables consistent alignment, which is crucial for downstream processing.
- **Blink Detection:** A CNN-based classifier processes cropped eye regions to determine whether the eyes are open or closed. A temporal buffer stores a sequence of predictions, and a blink is detected based on the transition pattern from 'open' to 'closed' and back to 'open.' This acts as a liveness check.
- **Anti-Spoofing Detection:** MiniFASNetV2, a lightweight CNN, classifies the facial region as real or spoofed. It detects texture inconsistencies by analysing depth and reflectance features that distinguish printed or screen-based spoofing attempts [12].
- **Mask Detection:** MobileNetV2, implemented using TensorFlow/Keras, is used to classify whether the detected face is masked or unmasked [13]. Its depthwise separable convolution layers provide computational efficiency suitable for lightweight execution.
- **Face Recognition:** InceptionResNetV1, pre-trained with ArcFace loss, generates a 512-dimensional embedding vector [14], [15]. A dual-embedding strategy is used, where both masked and unmasked embeddings are stored during registration [16]. During recognition, the system selects the appropriate embedding for comparison. Matching is performed using cosine similarity (1).

$$\text{similarity}(x, y) = 1 - \frac{x \cdot y}{\|x\| \|y\|} \quad (1)$$

- ArcFace enhances intra-class compactness and inter-class discrepancy by adding an angular margin penalty to the softmax loss. The ArcFace loss function is given by (2).

$$L = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s(\cos(\theta_{y_i} + m))}}{e^{s(\cos(\theta_{y_i} + m))} + \sum_{j \neq y_i} e^{s \cos(\theta_j)}} \quad (2)$$

where θ_{y_i} is the angle between the feature and the target class weight, m is the additive angular margin penalty, and s is a scaling factor.

Each component was developed and validated using publicly available datasets. Blink detection utilized open/closed eye datasets, mask detection was trained on RMFD and custom samples [13], [17], anti-spoofing was benchmarked on CASIA-FASD, and facial recognition relied on VGGFace2 with synthetic masking for robustness.

3.2 Workflow Execution

The authentication workflow begins with the user selecting Login, Logout, or Register. A camera image is captured and passed through each verification stage in sequence. First, a mask-status check is performed; then blink detection ensures liveness. If successful, the system proceeds to anti-spoofing, and finally to face verification. Access/logging occurs only if the face is verified against an enrolled template, anti-spoof = Real, and the preceding checks pass. When a mask is detected, the system selects the masked template for comparison; otherwise, it uses the unmasked template.

Attempts that fail any stage are marked invalid or routed to enrolment. We report only inference-relevant settings (input sizes, thresholds, gating policy, operating point, and evaluation protocol). User-interface details and screenshots are provided in Appendix A. Representative decision outcomes are provided in Appendix B.

A visual representation of this sequential process is presented in Figure 1, which outlines the interaction between components from image acquisition to timestamp logging.

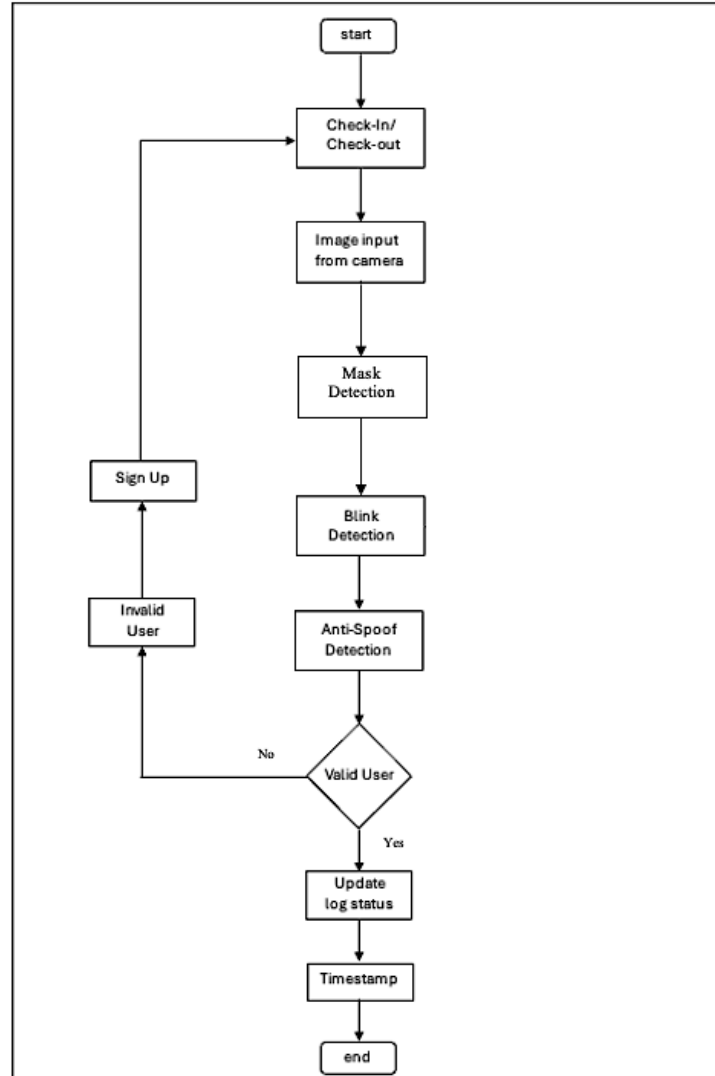


Figure 1. Flow Chart Diagram

3.3 Model Configurations, Operating Points and Inference Settings

We integrate off-the-shelf pretrained components and contribute a gated, mask-aware decision policy with dual-template routing and a transparent operating point (Blink ≥ 0.7 ; anti-spoof = Real; CosSim ≥ 0.75). Our focus is operational design and full-pipeline reporting rather than new backbones. Table 2 shows the module configuration and thresholds.

Hardware/runtime: Apple MacBook Air (M1), Apple M1 SoC (8-core CPU, integrated GPU via Metal), unified memory 8–16 GB; inference executed locally on macOS using CPU/Metal acceleration where applicable; batch size = 1.

Table 2. Module Configurations and Thresholds

Module	Pretrained model / files	Input and preprocessing	Key inference params	Decision rule / threshold	Notes
Face detection	RetinaFace (Python package retina-face) ; also, Caffe model for anti-spoof cropping: Widerface-RetinaFace.caffemodel + deploy.prototxt	Frame in BGR; library's default preprocessing	(Caffe detector used for cropping in anti-spoof) detector confidence = 0.60	-	We keep the highest-score face per frame.
Anti-spoofing	MiniFASNet/SE-ResNetweights (two models in models/anti_spoof_model/: 2_7_80x80..._FASNetV2.pth, 4_0_0_80x..._NetV1SE.pth	For each detected face, crop and resize per model spec (parsed from filename); softmax scores per model	Predictions from all .pth models are summed; label = $\text{argmax}([\text{Fake2D}, \text{Real}, \text{Fake3D}])$; score = $\text{prediction}[0][\text{label}]/2$	Live if label == 1 (Real)	No tuned threshold beyond the model's argmax; if not Real → "Spoof detected."
Blink / liveness	Blink classifier checkpoint: model_11_96_0.1256.t7 (loaded into BlinkModel) + dlib 68-landmark file shape_predictor_*.dat	Eye crops → grayscale, 24×24, Normalize mean=0.5, std=0.5	Smoothing buffer deque(maxlen=7); sampling every 0.5 s; timeout 10 s	Per frame: open_prob = $\text{softmax}[:,1]$, closed_prob = $\text{softmax}[:,0]$; Open if open_prob ≥ 0.7; Closed if closed_prob ≥ 0.7; Blink detected on Open→Closed transition; require ≥1 blink within timeout	Implements gating before authentication; uses MediaPipe Face Mesh for eye regions and head-pose checks below.
Head-pose straightness (gate)	PnP with 3D face model points (MediaPipe landmarks)	Uses camera intrinsics from frame size	-	yaw	≤ 35°
Mask detection	Keras mask_detector.model.h5 (MobileNetV2-style preprocessing)	Face crop → RGB → 224×224, preprocess_input(MobileNetV2)	Softmax [mask, no_mask]	Masked if and only if mask_prob > no_mask_prob	Files in models/mask_detection_model/also include res10_300x300...caffemodel + deploy.prototxt (not used in current inference path).
Face recognition (verification)	InceptionResNetV1 (ArcFace) checkpoint InceptionRes..._ArcFace.pt	Aligned face → 112×112; RGB; normalize (x-127.5)/128; L2-normalize embedding	Cosine similarity to enrolled template	Accept if cosine_similarity ≥ 0.75	RECOGNITION_THRESHOLD = 0.75 in code; embeddings normalized via F.normalize.
Decision fusion / operating rule	Heuristic AND rule (no learning)	-	-	Access granted if: (i) face straightness passes, and (ii) anti-spoof label == Real, and (iii) mask mode matches route (if applicable), and (iv) cosine ≥ 0.75, after a successful blink gate (≥1 blink)	Blink gate is executed first on /blink and before /login/register flows.

3.3.1 Enrolment and Templates

During registration, a face crop is captured and embedded; the L2-normalised embedding is stored as the template. At authentication, we compute the embedding for the live crop and compare via cosine similarity to the stored template. Integration strategy. We store two templates per user (masked and unmasked) at enrolment and select the matching template at inference based on the mask detector; verification then proceeds by cosine similarity at $\tau = 0.75$. This decision is gated by blink and anti-spoof outcomes; only samples with anti-spoof = Real and liveness passed proceed to verification.

3.3.2 Verification Scoring and Gated Decision Rule

Backbones are pretrained; we score at inference time only. Equation (3) shows the calculation of L2-normalization of embeddings while Equation (4) shows the Cosine similarity calculation.

$$\hat{\mathbf{f}} = \frac{\mathbf{f}}{\|\mathbf{f}\|_2} \quad (3)$$

$$s = \hat{\mathbf{f}}_q^\top \hat{\mathbf{f}}_t \quad (4)$$

Gated accept/reject rule (fixed OP) is shown at Equation (5). ArcFace loss is used as the pretrained verification loss [12].

$$\text{Accept} = 1 \left[\underset{\text{masked} \Rightarrow \text{masked template}}{\text{Mask-route}} \wedge (\text{Blink} \geq 0.7) \wedge (\text{Spoof} = \text{Real}) \wedge (s \geq 0.75) \right] \quad (5)$$

3.4 Datasets and Evaluation Protocol (Public datasets; notebook execution)

We did not collect new human data. All evaluations were performed offline in a Jupyter notebook using public datasets appropriate to each module, as shown in Table 3.

Table 3. Dataset Module Mapping

Module	Public dataset(s)	Task and labels	Split protocol	Notes
Anti-spoofing	CASIA-FASD	Live vs spoof (print/replay/2D)	Official/standard test split	Report spoof detection accuracy; no fine-tuning.
Blink / liveness	Open/Closed-Eye(CEW/OCEW or equivalent)	Open vs closed; sequence-level blink = open $\rightarrow$ closed	Subject-disjoint val/test	Used to set blink operating point; no user data collected.
Mask detection	RMFD ($\pm$ small, curated balance set, if used)	Mask vs no-mask	Subject-disjoint val/test	MobileNetV2 preprocessing at 224 $\times$ 224.
Face recognition (masked)	VGGFace2 (synthetic masked subset)	Same-ID vs different-ID (cosine verification)	Official pair list / subject-disjoint pairs	Used to choose τ ; no fine-tuning.

Backbones were not fine-tuned. For each task, we used identity-disjoint validation/test splits; operating points (e.g., cosine τ , blink and spoof decisions) were chosen on the validation split and then applied unchanged to the test split. End-to-end pipeline on public data. To assess integrated behaviour without new data collection, we executed the full pipeline sequentially on samples drawn from the public sets:

- Detection $\rightarrow$ crop (RetinaFace defaults)
- Anti-spoof ensemble (accept only when class = Real)
- Mask check (mask vs no-mask)
- Verification (cosine similarity ≥ 0.75) against the enrolled template of the claimed identity.

For blink gating, we detect blink using an eye-state classifier (open vs closed) on still images. If a short frame sequence is available, a blink is counted when an open frame is followed by a closed frame before verification. If only single images are available, the gate is satisfied if any captured still is classified as closed-eye; we report eye-state accuracy separately and label these runs as “blink (still-image)”. All conditions are those of the original public datasets; we did not alter or augment them. Our MacBook Air (M1) was used only to measure runtime (latency) and does not affect dataset labels or outcomes. We executed the full pipeline offline on public datasets. When short sequences were available, the blink gate was enforced; where only single images existed, we used blink (still-image). At the fixed operating point (cosine $\tau = 0.75$; anti-spoof accept = Real), we observe: [insert Acc / FAR / FRR / EER for the integrated run(s)]. Table 4 presents the operating point and evaluation protocol.

Table 4. Operating Point and Evaluation Protocol

Item	Specification
Operating point (OP)	Blink ≥ 0.7 ; Anti-spoof = Real (threshold 0.5); Face verification accept if CosSim ≥ 0.75
Pipeline order	Mask $\rightarrow$ Blink $\rightarrow$ Anti-spoof $\rightarrow$ Face verification (access/logging only if all gates pass)
Protocol	Identity-disjoint hold-out; choose OP on validation identities; apply unchanged to held-out test identities
Module $\leftrightarrow$ Dataset	Anti-spoof: CASIA-FASD (live vs spoof) $\cdot$ Blink: open/closed-eye dataset (still images; see note) $\cdot$ Mask: RMFD (mask vs no-mask) $\cdot$ Verification: masked subset / pairs from VGGFace2 (cosine)
Preprocessing	RetinaFace default crop; mask classifier at 224×224 ; embeddings L2-normalised; cosine similarity
Backbones	Pretrained only; no fine-tuning (RetinaFace, MobileNetV2, MiniFASNet, InceptionResNetV1 + ArcFace)
Blink gate note	For stills: eye-state classifier (open/closed); gate satisfied when a closed-eye is detected among captured frames. For short sequences: blink = open $\rightarrow$ closed before verification
Execution note	Offline notebook execution on public datasets; MacBook Air (M1) used only for latency measurements

3.5 Identity-Disjoint Hold-Out Evaluation (integration-only)

When integrated, the full authentication pipeline achieved an overall accuracy of 88%. Access is granted only when a user passes all stages (blink, spoof, mask, and recognition). The combined system enhances protection against common attack vectors while preserving usability. A detailed breakdown of the pipeline’s public-dataset, notebook-execution metrics is presented in Table 6. We evaluated the pipeline under an identity-disjoint hold-out protocol: operating points were selected on a validation split and then applied unchanged to a separate test split; all figures in Table 6 corresponds to the fixed operating point (Blink ≥ 0.7 , anti-spoof = Real, CosSim ≥ 0.75) applied to held-out identities. This indicates that the chosen operating point is stable under the public-dataset protocol, supporting the generalization claims for our integration-only pipeline.

4. RESULTS AND DISCUSSIONS

To situate our results in the broader literature, we benchmark against representative systems in Table 2. We focus on scientific evaluation (operating point, datasets, identity-disjoint hold-out) and report UI details only in Appendix A. Representative decision outcomes are provided in Appendix B. To evaluate the performance of each module in the proposed Healthcare Authentication System, we conducted experiments using benchmark datasets. The key metrics used include precision, recall, F1-score, and overall accuracy. These results validate the effectiveness of the system components in handling masked face recognition, spoof detection, and liveness verification. The accuracy results of each module, tested using their respective benchmark datasets, are summarized in Table 5.

Table 5. Accuracy of Individual Biometric Modules

Module	Dataset Used	Accuracy
Blink Detection	Open/Closed Eye Datasets	91%
Spoof Detection	CASIA-FASD	85%
Mask Detection	RMFD + custom images	97%
Face Recognition	VGGFace2 (masked)	96%
Full Pipeline	Combined evaluation	88%

The blink detection model achieved an accuracy of 91% using the open/closed eye dataset. Precision and recall scores were balanced at 0.91 for both classes (“Open” and “Closed”), with a macro and weighted F1-score of 0.91, indicating that the model reliably distinguished between live and spoof attempts involving static eyes. The MiniFASNetV2-based anti-spoofing model achieved an overall accuracy of 85% on the CASIA-FASD dataset. The model achieved precision of 0.78 (fake) and 0.95 (real), while the recall was 0.96 (fake) and 0.73 (real). This shows a strong ability to block spoof attempts, though some real users were occasionally misclassified. The F1-scores of 0.86 (fake) and 0.83 (real) show balanced predictive power, and the confusion matrix indicates robust spoof rejection with occasional false negatives. The use of MiniFASNet ensures lightweight performance on constrained devices, similar to other mobile-efficient models such as ShuffleNet [20].

The mask detection module trained on RMFD, and custom images achieved 96% accuracy on a test set of over 8,000 images. Both classes (“Mask” and “No Mask”) showed precision and recall above 0.95, with F1-scores of 0.96–0.97. The confusion matrix reflects very few misclassifications, confirming the reliability of mask presence detection in lightweight conditions.

The InceptionResNetV1 model trained with ArcFace loss demonstrated 96% accuracy on the masked face dataset. Precision ranged from 0.93 to 0.99, and recall from 0.93 to 0.99 across both classes, with F1-scores of 0.96. This high consistency validates the effectiveness of the dual-embedding strategy in masked/unmasked scenarios. The distribution graph also confirms the separation of genuine and imposter scores.

The full authentication pipeline achieved 88% accuracy, granting access only when all stages are passed: mask check → blink (liveness) → anti-spoof → face verification. This integrated design strengthens protection against common attacks while maintaining usability. Table 6 presents the pipeline’s public-dataset, notebook-execution metrics, evaluated under an identity-disjoint hold-out protocol. All reported figures use the same fixed operating point (Blink ≥ 0.7 , anti-spoof = Real, CosSim ≥ 0.75), applied unchanged to held-out identities. The identity-disjoint evaluation confirms the stability of this operating point and supports the generalisation of our integration-only pipeline.

Although the system employs pretrained components, the integration strategy and gating policy—particularly the sequencing of blink and anti-spoof checks prior to verification and the use of dual templates for masked and unmasked faces—produce robust end-to-end performance (88%) at a conservative, low-FAR operating point on public-dataset inputs.

To evaluate each component under a consistent operating point, we consider FR-only, FR+mask, FR+spoof, FR+blink, and the full pipeline. FR-only exhibits higher FAR; adding mask improves template selection under occlusion, spoof reduces presentation-attack errors, and blink blocks static-photo attacks.

The full integration achieves a conservative, low-FAR operating point with 88% end-to-end accuracy (Table 6). Operating point. Blink ≥ 0.7 ; Anti-spoof = Real (0.5); CosSim ≥ 0.75 (identity-disjoint hold-out; see Table 2).

Table 6. Final Pipeline Evaluation Metrics

Metric	Value	Description
Total Attempts	76	Number of public-dataset attempts (offline, notebook-execution)
Correct Decisions	67	TP + TN at the fixed operating point
Final Accuracy	88%	Overall pipeline success rate
Imposter Precision	0.79	How often predicted imposters were truly imposters
Imposter Recall	1.00	Ability to catch all spoof attempts
Genuine Precision	1.00	How often predicted genuine users were correct
Genuine Recall	0.79	Ability to authenticate all real users
F1-Score (both classes)	0.88	Harmonic mean of precision and recall
ROC AUC (verification)	>0.95	Strong separation between genuine and imposter score distributions
Threshold Settings	Blink ≥ 0.7 Spoof ≥ 0.5 CosSim ≥ 0.75	Fixed operating point used across results

Operating point. Blink ≥ 0.7 ; Spoof = Real (0.5); CosSim ≥ 0.75 (identity-disjoint hold-out on public datasets; see Table 2). These results demonstrate consistent end-to-end behaviour under our public-dataset, notebook execution. At the fixed operating point (Blink ≥ 0.7 , Spoof ≥ 0.5 , CosSim ≥ 0.75), the full authentication pipeline achieves an overall accuracy of 88% (Table 6). On the final imposters-vs-genuine classification, the system attains precision = 0.79 / 1.00 and F1 = 0.88 (imposter / genuine), indicating a conservative, low-FAR operating point. Overall accuracy refers to the end-to-end pipeline evaluated at Blink ≥ 0.7 , Spoof ≥ 0.5 , and CosSim ≥ 0.75 with the identity-disjoint protocol. We report ablations under an identical, fixed operating point (Blink ≥ 0.7 ; Spoof = Real (0.5); CosSim ≥ 0.75); see Section 2.4 for external benchmarks and Table 2 for protocol details.

The baseline/ablation view (Table 7) clarifies contribution: FR-only exhibits higher FAR; adding mask routing improves matching under occlusion; adding anti-spoof removes spoof accepts; adding blink blocks static-photo attacks. The full gated pipeline thus achieves 88% with a conservative FAR profile. Qualitative panels in Appendix B illustrate these behaviours (blink pass/fail, spoof blocked, mask-aware template selection), aligning visual outcomes with the quantitative evidence. Overall, the findings justify the chosen OP for deployment settings where false accepts are high-risk, while the identified error modes (occlusion, pose, near-threshold scores) motivate our future work on fusion calibration, temporal liveness, and risk-aware thresholding.

Table 7. Baselines under Identical Protocol (Identity-disjoint Hold-out; Fixed OP).

Setup	Blink gate	Anti-spoof	Mask routing (dual template)	Verification threshold	Expected effect on FAR	Expected effect on Genuine Recall	Notes
FR-only	✗	✗	✗	CosSim ≥ 0.75	Higher	Higher (may accept more)	No defences; prone to spoof and static photo attacks
FR + Mask	✗	✗	✓	CosSim ≥ 0.75	Higher	Similar/Higher	Helps occlusion by using matched template
FR + Spoof	✗	✓	✗	CosSim ≥ 0.75	Lower	Lower (rejects some genuine)	Blocks presentation attacks
FR + Blink	✓	✗	✗	CosSim ≥ 0.75	Lower	Similar/Lower	Blocks static photo attacks
Full (ours)	✓	✓	✓	CosSim ≥ 0.75	Lowest (conservative OP)	Moderate (balanced)	Matches reported 88% end-to-end

At the fixed operating point (Blink ≥ 0.7 ; Spoof = Real (0.5); CosSim ≥ 0.75), the end-to-end pipeline favours low FAR while retaining acceptable genuine throughput: Impostor recall = 1.00 and Genuine recall = 0.79, yielding F1 = 0.88.

Cross-checking the module summaries (Tables 8 - 11) shows that anti-spoofing is deliberately conservative (minimising presentation-attack risk), and that most residual errors concentrate near the cosine threshold for masked/partially occluded or mild-pose faces.

Table 8. Anti-spoof (CASIA-FASD)

Precision	Recall	F1	Accuracy
0.87	0.85	0.84	0.85

Table 9. Face verification (masked subset)

Precision	Recall	F1	Accuracy
0.96	0.96	0.96	0.96

Table 10. Blink (open/closed)

Precision	Recall	F1	Accuracy
0.91	0.91	0.91	0.91

Table 11. Mask detection (RMFD)

Precision	Recall	F1	Accuracy
0.91	0.88	0.88	0.88

The fully integrated pipeline was evaluated on public-dataset inputs in an offline notebook execution, achieving an overall accuracy of 88% at the fixed operating point (Blink ≥ 0.7 , anti-spoof = Real, CosSim ≥ 0.75). In one representative public-dataset subset, the model correctly identified all imposter attempts (11/11) and 78.6% of genuine users (11/14); the corresponding confusion matrix is shown in Figure 2. This reflects a conservative operating point (low FAR) while maintaining reliable identification of genuine users. Figure 3 shows the precision, recall and F1 score for imposter and genuine users.

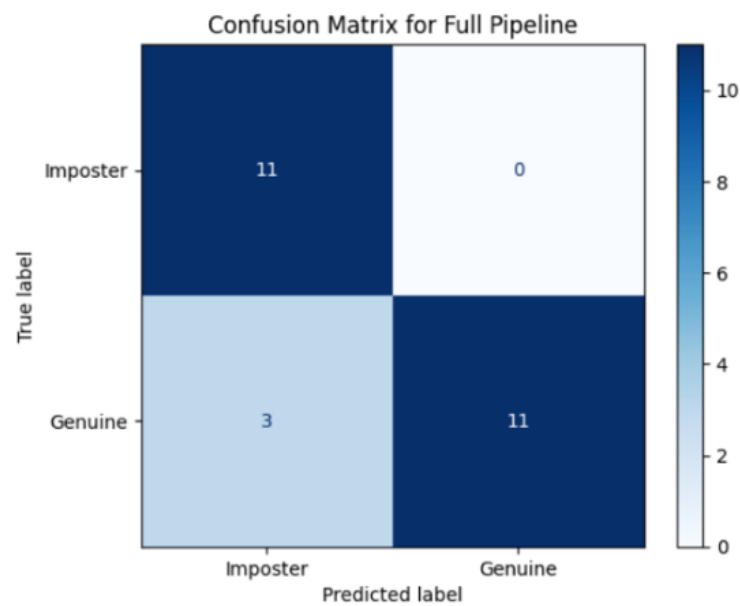


Figure 2. Confusion Matrix for a Representative Public-dataset Subset

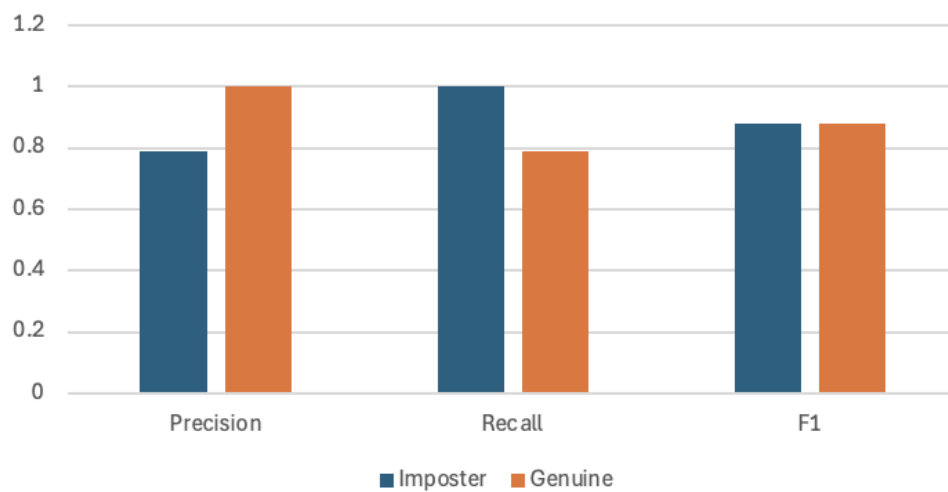


Figure 3. Precision / Recall / F1 for Imposter and Genuine

5. CONCLUSION

Main findings. Under the public-dataset, identity-disjoint hold-out protocol and a fixed operating point ($\text{Blink} \geq 0.7$; $\text{Anti-spoof} = \text{Real}$ (0.5); $\text{CosSim} \geq 0.75$), the integrated pipeline achieves 88% end-to-end accuracy with a conservative (low FAR) operating profile. Class-wise behaviour shows Impostor recall = 1.00 (all 11 spoofs rejected) and Genuine recall = 0.79 (11/14 accepted), consistent with the error distribution observed in the confusion matrix (Figure 2). Module-level summaries confirm that each component performs robustly at the selected operating point, and the baseline comparison clarifies the incremental value of mask routing, anti-spoof, and blink gating.

Contributions: (i) An integration-first pipeline tailored to healthcare: mask-aware dual-template verification gated by blink liveness and anti-spoof; (ii) a transparent operating point and identity-disjoint evaluation on public datasets, reported end-to-end (not only module metrics); and (iii) a lightweight, deployable implementation that makes explicit the security–usability trade-offs through fixed thresholds and clear reporting.

Practical implications: The gating policy prioritizes low false accepts—appropriate for clinical access control—while maintaining acceptable genuine throughput. The dual-template strategy improves recognition under mask usage common in hospitals, and the end-to-end reporting provides implementers with auditable thresholds they can adopt or tune for their risk tolerance.

Limitations: We use pretrained backbones without fine-tuning, evaluate offline on public datasets rather than live clinical trials, and operate with fixed thresholds rather than learned fusion/calibration. Residual errors concentrate near the cosine threshold for masked/partially occluded or mild pose faces and from occasional blink misses under head tilt.

Future work (high-level):(1) Fusion and calibration: replace heuristic AND-gating with a lightweight learned fusion calibrated at target FAR, following recent advances in adaptive multimodal biometric fusion [21]. (2) Broader evaluation: larger, multi-dataset identity-disjoint and cross-dataset transfer to quantify domain shift. (3) Temporal liveness: short-sequence blink/micro-motion to reduce misses in low-light/pose. (4) Sensors abd deployment: depth/IR where available; on-device optimization (Apple Silicon / embedded) with latency and memory reporting. (5) Security and fairness: template protection and audit logging; stratified FAR/FRR across demographic and environmental factors. In addition, privacy-preserving and continuous biometric authentication approaches, such as federated learning and privacy-enhancing techniques, could be investigated to support secure deployment in healthcare environments [22].

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers for their valuable comments.

FUNDING STATEMENT

The authors received no funding from any party for the research and publication of this article.

AUTHOR CONTRIBUTIONS

Caleb Yeo: Conceptualization, Methodology, Validation, Writing – Original Draft Preparation;
Siew-Chin Chong: Project Administration, Data Curation, Supervision, Writing – Review and Editing;
Kian-Ming Lim: Methodology, Validation;
Lee-Ying Chong: Project Administration, Writing – Review and Editing.

CONFLICT OF INTERESTS

No conflict of interest was disclosed.

ETHICS STATEMENTS

No new human participants were recruited and no new personal data were collected. The study used publicly available datasets.

DATA AVAILABILITY

The datasets used in this study are publicly available. CASIA-FASD, RMFD, and VGGFace2 can be accessed from their respective public repositories at <http://biometrics.idealtest.org>, <https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset>, and https://www.robots.ox.ac.uk/~vgg/data/vgg_face2/.

REFERENCES

- [1] Z. A. Bahman, "Biometrics in healthcare as a technological contribution to mitigate patient identification errors: A review," *International Journal of Development Research*, pp. 64930–64937, 2024, doi: 10.37118/ijdr.27834.02.2024.
- [2] Pew Research Center, "Pew Research Center," 2024. [Online]. Available: <https://www.pewresearch.org/>
- [3] Y. Atoum, Y. Liu, A. Jourabloo, and X. Liu, "Face anti-spoofing using patch and depth-based CNNs," *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pp. 319–328, 2017, doi: 10.1109/btas.2017.8272713.
- [4] D. N. Nishanth and G. Mallikarjuna Rao, "Liveness detection based on human eye blinking for photo attacks," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 1, pp. 4074–4077, 2019, doi: 10.35940/ijeat.a1314.109119.
- [5] M. Ngan, P. Grother, and K. Hanaoka, "Ongoing face recognition vendor test (FRVT) part 6A," *NIST Interagency Report*, 2020, doi: 10.6028/NIST.IR.8311.
- [6] D. Montero, M. Nieto, P. Leskovsky, and N. Aginako, "Boosting masked face recognition with multi-task ArcFace," *arXiv preprint*, 2022, doi: 10.1109/SITIS57111.2022.00042.
- [7] S. Shuhui and M. Xiaochen, "Dual-proxy modeling for masked face recognition," *Procedia Computer Science*, vol. 208, pp. 145–151, 2022, doi: 10.1016/j.procs.2022.10.022.
- [8] Y. Himeur *et al.*, "Face mask detection in smart cities using deep and transfer learning: Lessons learned from the COVID-19 pandemic," *Systems*, vol. 11, no. 2, pp. 107, 2023, doi: 10.3390/systems11020107.
- [9] M. I. Razzak, M. Imran, and G. Xu, "Big data analytics for preventive medicine," *Neural Computing and Applications*, vol. 32, no. 9, pp. 4417–4451, 2019, doi: 10.1007/s00521-019-04095-y.
- [10] A. Holzinger, C. Biemann, C. S. Pattichis, and D. B. Kell, "What do we need to build explainable AI systems for the medical domain?" *arXiv preprint arXiv:1712.09923*, 2017. [Online]. Available: <https://arxiv.org/abs/1712.09923>
- [11] Z. Qin *et al.*, "A survey of identity recognition via data fusion and feature learning," *Information Fusion*, 2022, doi: 10.1016/j.inffus.2022.10.032.
- [12] Z. Yu *et al.*, "Searching central difference convolutional networks for face anti-spoofing," *arXiv preprint*, 2020, doi: 10.1109/CVPR42600.2020.00534.
- [13] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. -C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.

- [14] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the impact of residual connections on learning," in *AAAI'17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, pp. 4278–4284, 2017.
- [15] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, pp. 4685–4694, 2019, doi: 10.1109/CVPR.2019.00482.
- [16] Z. Wang, B. Huang, G. Wang, P. Yi, and K. Jiang, "Masked face recognition dataset and application," in *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 5, no. 2, pp. 298–304, April 2023, doi: 10.1109/TBIOM.2023.3242085.
- [17] A. Cabani, K. Hammoudi, H. Benhabiles, and M. Melkemi, "MaskedFace-Net – A dataset of correctly/incorrectly masked face images in the context of COVID-19," *Smart Health*, vol. 19, pp. 100144, 2021, doi: 10.1016/j.smhl.2020.100144.
- [18] N. D. Lane *et al.*, "DeepX: A Software Accelerator for Low-Power Deep Learning Inference on Mobile Devices," *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, Vienna, Austria, pp. 1–12, 2016, doi: 10.1109/IPSN.2016.7460664.
- [19] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, "RetinaFace: Single-shot multi-level face localisation in the wild," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 5202–5211, 2020, doi: 10.1109/CVPR42600.2020.00525.
- [20] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 6848–6856, 2018, doi: 10.1109/CVPR.2018.00716.
- [21] P. Chitrapu, M. K. Morampudi, and H. K. Kalluri, "A secure and explainable multimodal biometric system using trust adaptive fusion for face and fingerprint," *Scientific Reports*, vol. 16, p. 14244, 2026, doi: 10.1038/s41598-026-43252-x.
- [22] A. Al Abdulwahid, "Privacy-aware continuous federated biometric authentication (PACFBA) using federated deep learning: A novel usable security framework," *Scientific Reports*, 2026, doi: 10.1038/s41598-026-55046-2.

BIOGRAPHIES OF AUTHORS

	Caleb Yeo is a final-year Information Technology student at Multimedia University, specializing in Security Technology. Fluent in English, Mandarin, and Malay, he excels in Java, C++, Python, and Web Development. His projects include an iNVENTX 2025 Silver Award-winning healthcare authentication system featuring liveness verification and mask detection. Passionate about cybersecurity and AI, Caleb aims to develop secure, lightweight tech solutions. He can be contacted at email: caleb3721@gmail.com.
	Siew-Chin Chong, an IEEE Senior Member, holds a B.IT, M.Sc, and Ph.D. in Information Technology from Multimedia University. She is currently Deputy Dean of Student Experience and Alumni at the Faculty of Information Science and Technology, MMU. Her research focuses on machine learning, biometric security, and mobile app development, with numerous publications. She has also served on journal editorial boards and as Technical Chair for international conferences. Contact: chong.siew.chin@mmu.edu.my.

	Kian-Ming Lim, an Associate Professor at the University of Nottingham Ningbo China and IEEE Senior Member, was recognized among the World's Top 2% Scientists by Stanford University in 2025. His research focuses on machine learning, deep learning, computer vision, few-shot learning, federated learning, generative AI, and NLP. He can be contacted at: Kian-Ming.Lim@nottingham.edu.cn.
	Lee-Ying Chong received her B.IT (Hons.) in Information System Engineering (2003), M.Sc. in Information Technology (2007), and Ph.D. in Information Technology (2018) from Multimedia University. Her research interests include biometric authentication, computer vision, and machine learning. She has been an IEEE Senior Member since 2013. Currently, she is a senior lecturer at the Faculty of Information Science and Technology, Multimedia University, Malaysia. She can be contacted at: lychong@mmu.edu.my.

APPENDIX A

This appendix summarises the user interface for reproducibility. The main paper reports only inference-relevant settings (thresholds, gating, operating point) and evaluation protocol. The home page in Figure A.1 presents a centred logo, an Admin button, and three actions—Login, Logout, and Register. The Admin button opens a password gate to view the Punch Log and User List. Login/Logout start the authentication flow; Register opens enrolment.

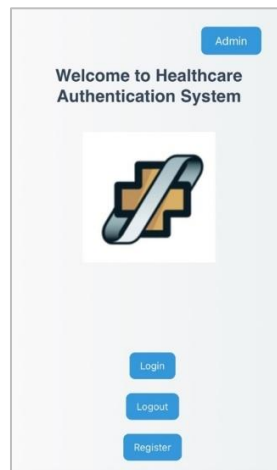


Figure A.1. Home Page

After password entry, the Admin page in Figure A.2. shows two tables: Punch Log (login/logout records) and User List (with remove action).

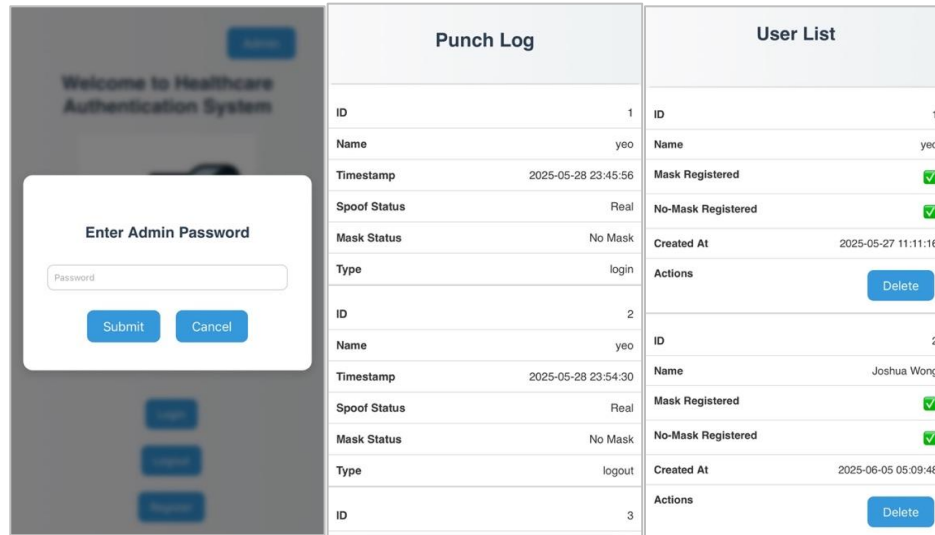


Figure A.2. Admin Page

All flows (Login, Logout, Register) start with blink verification. The user faces the camera; misalignment triggers an on-screen warning. After a three-second countdown, a blink must be detected to proceed. If blink fails or the face deviates, the system displays “Blink detection failed” and returns to the home page. Figure A.3. shows the blink detection process.

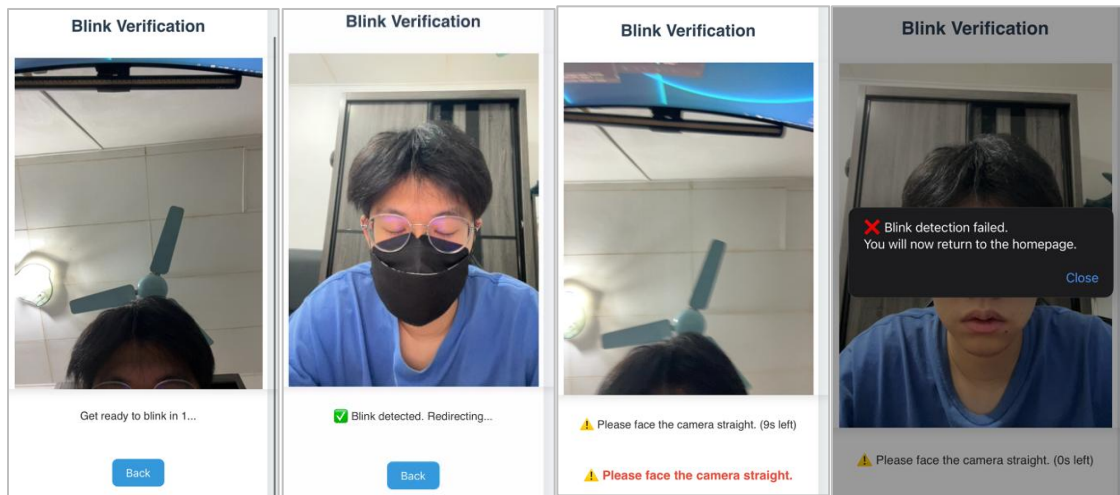


Figure A.3. Blink Detection

The pipeline executes blink → anti-spoof → mask check → face verification. Access/logging occurs only if anti-spoof = Real and the embedding matches an enrolled template. Success messages (“Welcome [Name]”, “Good-bye [Name]”) are shown on acceptance; “Spoof detected” prevents logging. Figure A.4. demonstrates the login and logout page for the spoof detection.

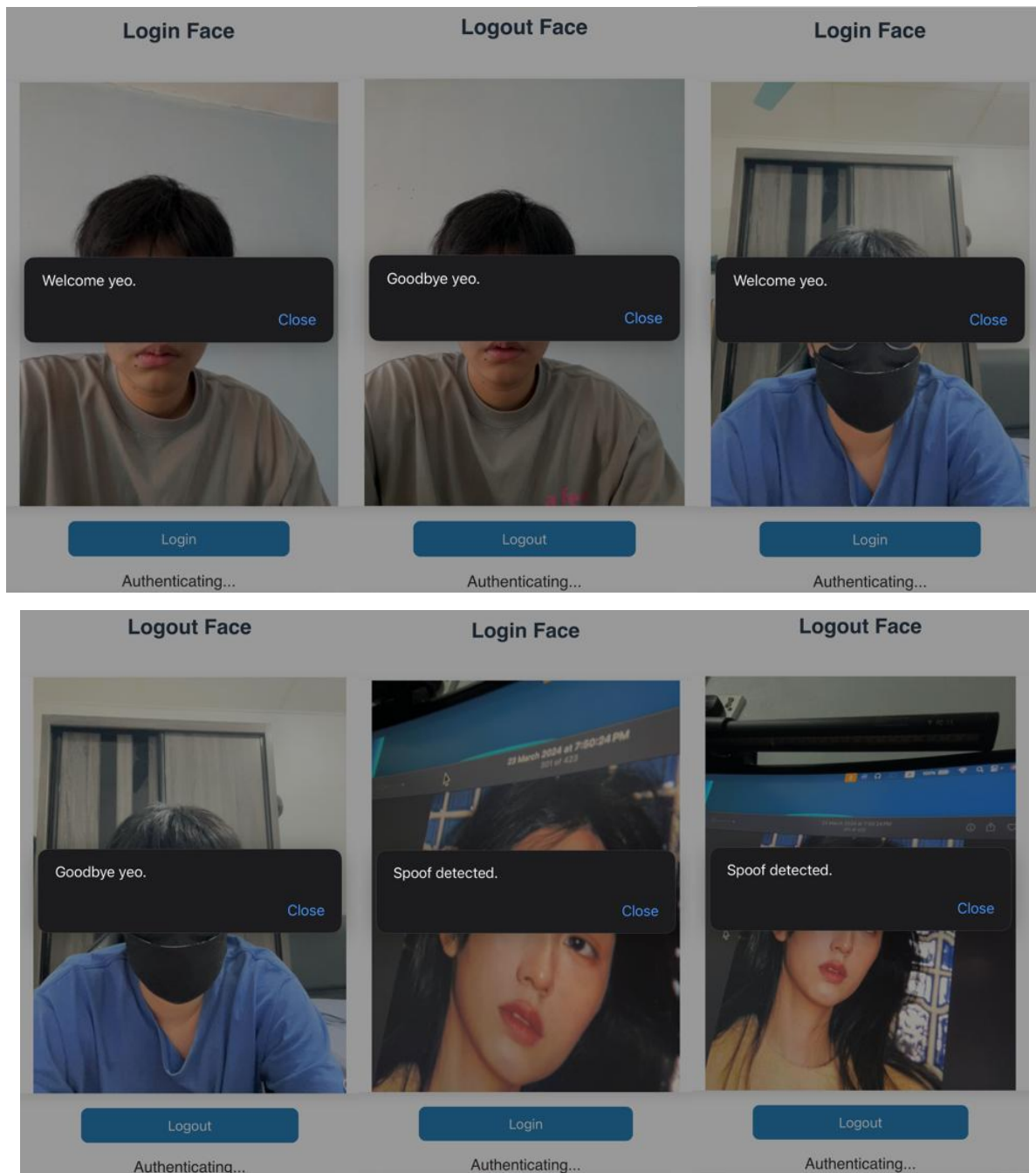
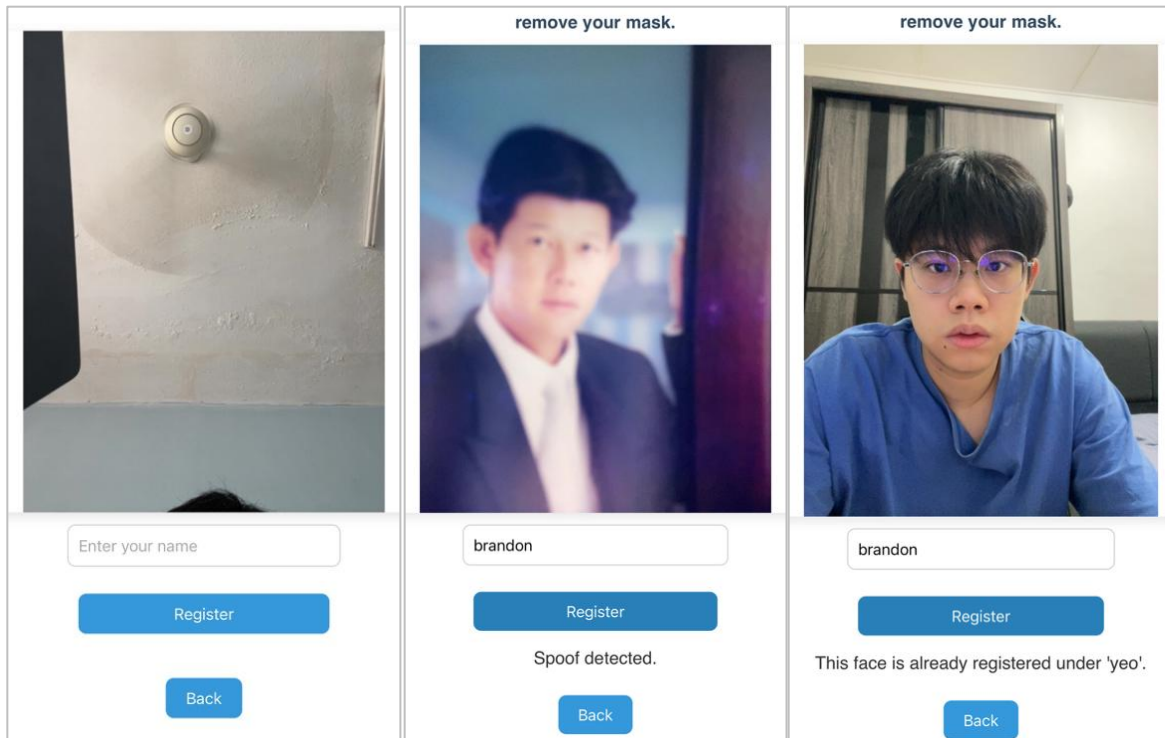


Figure A.4. Login and Logout Page



Enrolment performs: (i) duplicate checks (name/embedding); (ii) anti-spoof (reject if fake); (iii) mask-status capture for unmasked and masked templates. If a required mask condition is not met, an error prompts retake. On success, both templates are stored and the user is returned to the home page, as shown in Figure A.5.

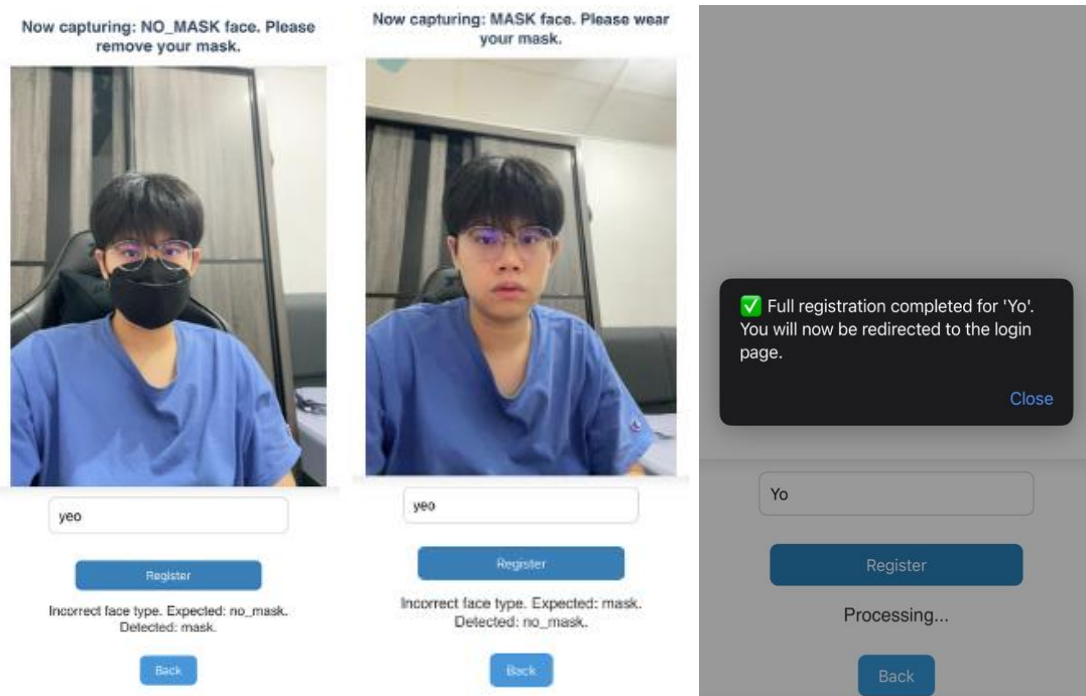
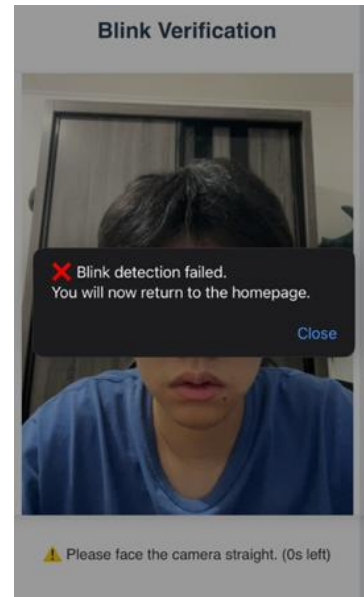
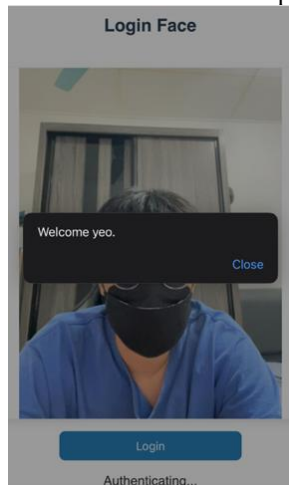
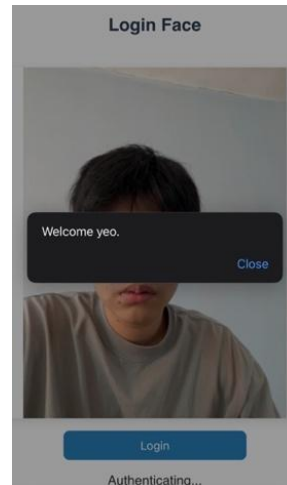
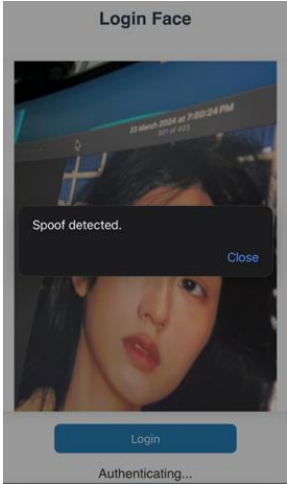
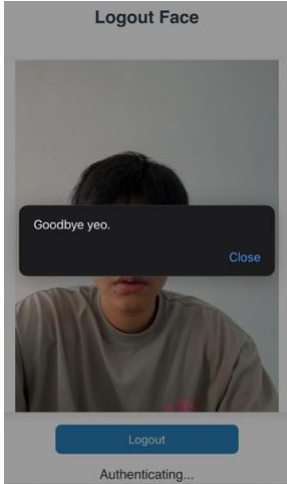


Figure A.5. Register Page

APPENDIX B – Qualitative Outcomes**a. Blink / Liveness****i. Blink pass → “Blink detected: proceeding...”****ii. Blink fail → “Blink detection failed”****b. Mask routing (dual templates)****i. Mask = YES → “masked template selected”****ii. Mask = NO → “unmasked template selected”**

c. Anti-spoof

i. Spoof detected → “Spoof detected”  The screenshot shows the 'Login Face' interface. At the top, it says 'Login Face'. Below that is a video feed of a person's face. A black overlay box in the center of the video feed contains the text 'Spoof detected.' and a 'Close' button. At the bottom of the interface, there is a blue 'Login' button and the text 'Authenticating...'.	ii. Real face → “Goodbye username”  The screenshot shows the 'Logout Face' interface. At the top, it says 'Logout Face'. Below that is a video feed of a person's face. A black overlay box in the center of the video feed contains the text 'Goodbye yeo.' and a 'Close' button. At the bottom of the interface, there is a blue 'Logout' button and the text 'Authenticating...'.
---	--