

GeoSecure-FR: A Geofenced, Occlusion-Aware Face Recognition Framework for Proximity-Based Attendance and Access Verification

R. Pradeep Kumar¹, M. Indumathy²

^{1,2}Department of Computer Science and Engineering, SRM Institute of Science and Technology
Vadapalani, Chennai, Tamil Nadu, India

Abstract

Attendance systems that match a face against a gallery answer the wrong question. They ask whether an image resembles an enrolled person, not whether a live, visible, enrolled person is at the venue. This paper presents GeoSecure-FR as a rejection cascade for that problem. A phone sends its GPS fix. The method arms the camera only when the Haversine distance to the venue centre is inside a configured radius plus a GPS slack. A captured frame must then clear a visibility check. The serving liveness gate samples a short frame window. While a recapture remains, a missing blink stops the attempt before recognition. The liveness flag is the disjunction of a detected blink and a MobileNetV2 score of at least 0.80, and only the last window is accepted or rejected by that flag. The first failure stops the chain, and the log stores the failing stage. On the AT&T face set, the rank-1 result for the matcher alone is 1.000 for ArcFace against 0.865 for an Eigenfaces baseline under the same protocol. That comparison does not pass the frame through the geofence, the visibility gate, or liveness. A lower-face blackout rule, added after three weaker rules failed, rejects 200 of 200 synthetically masked probes and none of 400 clean frames. Retraining the liveness model on a CelebA-Spoof subset moves two screen-replay scores from 0.986 and 0.968, both false accepts, to 0.035 and 0.121 at a live threshold of 0.80. CPU end-to-end latency remains about 4–7 s, above a 3 s target. These figures show the sample and the condition behind each number. They do not show a presentation-attack rate, a real-mask rate, room-level location, or a latency below 3 s.

Keywords: Face Recognition, Geofencing, Occlusion, Liveness Detection, Attendance, Access Control

1. Introduction

Manual roll calls, RFID cards, and QR scans record that a token was presented. They do not record that a particular person was standing in the room. Face recognition removes the token, and published classroom systems can replace a roll call with a camera [1, 2]. Three failures remain. A printed photograph or a phone screen can satisfy a matcher that never tests liveness [3]. A mask, a hand, or a heavy shadow pushes the embedding off the enrolled identity, and a closed-set matcher may still return the nearest name. Nothing in the pipeline verifies that the image was captured at the venue. A face sent from another campus is treated as presence.

Jamdar et al. combine a geofence with a MobileNetV2 liveness model for attendance [4]. That pairing blocks two of the three failures. An occluded face is still passed to the matcher, and a denial is not tied to

the stage that produced it. Mask-robust recognition takes the other stance: it tries to identify a covered face instead of asking for another frame. GeoSecure-FR keeps recognition for the last stage and refuses the attempt at an earlier stage.

The implemented system consists of a Flask service and a phone or browser client. Enrollment is a command-line step that stores a mean ArcFace embedding. Check-in is a location test. Verification runs only after that test passes. Each attempt is recorded as accepted, denied, or spoof-flagged, with the gate that stopped it.

This paper makes three claims, and each claim is limited to the evidence in Section 4.

1. Ordering geofence, visibility, liveness, and open-set matching, and withholding the camera until the fence passes, separates remote submission from an at-venue attempt. The audit record names the failing gate.
2. A visibility gate that rejects and asks for a new frame avoids matching a covered face. On synthetic lower-half masks over the AT&T set, the adopted rule catches every masked probe without rejecting a clean frame. Earlier variants of the same gate do not.
3. An Eigenfaces baseline [5] and a raw ArcFace matcher [6] show what the gallery comparison is worth on clean faces, and how badly that comparison behaves once the lower face is covered and no gate is present.

The paper does not claim a new embedding, a trained occlusion classifier, or a presentation-attack error rate. Those items are stated as open in Section 5.

2. Related Work

Classroom attendance papers typically detect a face and assign a label. Boe et al. use a histogram-of-oriented-gradients detector with a support-vector machine and a convolutional identifier, and compare the result with QR attendance [1]. Kumar et al. combine the same detection family with a convolutional encoding and a temperature-and-sanitizer side check [2]. Both confirm appearance. Neither confirms that the person is inside a place, nor do they refuse a frame because the face is covered.

Jamdar et al. are the closest published design in the set reviewed for this project [4]. Their system uses facial recognition, a MobileNetV2 liveness model trained on a custom set, and geofencing, specifically to reduce false attendance. The occlusion case is left to the matcher. GeoSecure-FR adds an explicit visibility stage in front of liveness and recognition, and it records denials by stage rather than as a single negative outcome.

The recognition and detection pieces are used as published. ArcFace adds an additive angular margin so that embeddings of the same person tighten and embeddings of different people separate [6]. FaceNet is the earlier triplet-loss embedding and is not the matcher in this build [7]. MTCNN supplies the face box and landmarks used by the visibility rules [8]. Eigenfaces, a principal-component nearest neighbour on aligned grayscale faces, is the baseline [5]. For presentation attacks, Yu et al. review the move from hand-crafted texture cues to deep detectors and the gap between intra-dataset scores and attacks the detector was not trained on [3]. CelebA-Spoof is the replay-oriented set used for the second liveness training run: 625,537 images of 10,177 subjects, with print, paper-cut, replay, and three-dimensional spoofs [9]. The liveness network itself is MobileNetV2 [10].

3. Method

Fig. 1 is the decision order for one attempt. The server does not run detection, liveness, or matching on an attempt that fails the fence. A later gate is not consulted after an earlier denial.

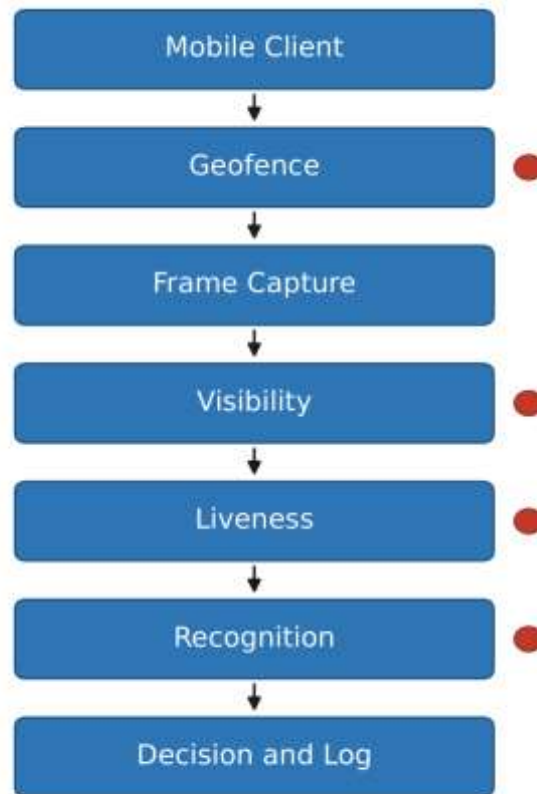


Fig. 1. Decision order for one attempt. Visibility is the composite visibility score and the lower-face rule. Liveness is a detected blink or a MobileNetV2 score of at least 0.80. A red mark is a stage that can stop the attempt, and a later stage is not run after an earlier denial.

A gray stage on the dashboard means the stage was never reached.

3.1 Geofence

Let ϕ be latitude and λ longitude in radians, and let R be the Earth radius used by the implementation, 6371 km. For a device fix (ϕ_1, λ_1) and a venue centre (ϕ_2, λ_2) ,

$$\begin{aligned} a &= \sin^2(\Delta\phi/2) + \cos \phi_1 \cos \phi_2 \sin^2(\Delta\lambda/2), \\ d &= 2R \operatorname{atan2}(\sqrt{a}, \sqrt{1-a}). \end{aligned} \quad (1)$$

where d is the great-circle distance in kilometres, and the attempt is inside the fence when $d \leq r + s$.

The operating values are $r = 300$ m and $s = 20$ m, so the cutoff used in tests is 320 m.

The slack absorbs ordinary GPS error near the boundary.

Indoor error of tens of metres remains larger than a classroom wall, so the fence is a building-scale test, not a room test.

If the client sends no fix, the gate fails closed and the camera is not armed.

The fence centre and radius live in configuration.

The coordinates in a request are the claimed fix for that attempt.

Moving the fence to imitate an off-site user is a different experiment from submitting a distant fix.

3.2 Visibility

The as-built gate is a composite score, not a trained four-class network.

MTCNN supplies a box, a confidence, and landmarks [8].

The score also uses face area, a Haar eye check, a mouth-region flag, a yaw proxy, and the lower 45% of the box.

That lower band contributes its mean intensity and its Laplacian variance.

A frame passes when the composite visibility is at least 0.70.

The client may recapture up to three times.

The denial reason is stored.

Four rules were compared on the same probes, and only the last two are part of the adopted gate.

A hard mouth-region flag, which had previously only subtracted 0.25 from the score, did not catch the masked frames that the baseline already missed.

Upscaling the native AT&T frame, or applying histogram equalization before MTCNN, raised the clean false-reject count and lowered the masked catch count.

The rule that closed the gap rejects a box when the lower-band mean intensity is at most 47.319 and the lower-band Laplacian variance is at most 176.698.

A Haar fallback covers the one clean frame MTCNN missed.

Sunglasses in the live demo often pass this gate and fail later, at liveness.

The rule is aimed at a covered lower face.

3.3 Liveness

The check that runs on POST /verify is an eye-blink test over a sampled window, followed by a texture score on one frame of that window.

MediaPipe Face Mesh supplies the landmarks.

If that API is absent, the same module loads `models/face_landmarker.task`.

For one eye the aspect ratio is

$$\text{EAR} = (\|p_2 - p_6\| + \|p_3 - p_5\|) / (2 \|p_1 - p_4\|), \quad (2)$$

where $p_1, \dots, p_6$ are the six-point eye contour of Soukupová and Čech [11].

The frame value is the mean of the two eyes.

A blink is the sequence open, closed, open: EAR at least 0.20, then at most 0.15 for two consecutive valid frames, then at least 0.20 again.

At least eight frames in the window must yield landmarks.

A missing frame breaks a closed run, so one low sample is not a blink.

The capture page sends about 13 frames from a 1.8 s window and tells the user to blink once.

A window is live when a blink is detected or $P(\text{real}) \geq 0.80$.

A detected blink is not rejected because the texture score is low.

MobileNetV2 produces that score on the highest-EAR frame, using the 0.20 crop margin from training [10].

While a recapture remains, a window with no blink is returned for another capture and is not passed to recognition.

The server allows two recaptures.

A texture score of at least 0.80 does not accept those earlier windows, because the liveness flag is not read yet.

Only the last window is accepted or rejected by the flag. On that window, $P(\text{real}) \geq 0.80$ continues to recognition even when no blink was observed.

If the weight file is absent, the texture term falls back to Laplacian variance, local texture variance, and hue standard deviation at a threshold of 0.50, and the log names that path.

The replay numbers in Section 4 are texture scores on still images.

They are not counts of blink trials.

Two training defects were found and corrected before the numbers in Section 4.

A random split of the NUAA print set placed the same subject on both sides, which produced a perfect and meaningless validation score.

Identity-disjoint splits removed that leak.

Separately, the pipeline cropped the detector box with no margin, while training crops used a margin of 0.20.

The same genuine photograph scored 0.41 in the pipeline and 0.976 as an isolated crop.

The pipeline crop now uses the training margin.

NUAA does not contain screen replay.

A balanced subset of 20,000 CelebA-Spoof images, with about half of the spoof class drawn from replay, was used for a two-phase fine-tune: twelve epochs with the backbone frozen, then eight epochs of fine-tuning [9].

3.4 Recognition and Log

The matcher is ArcFace, called through DeepFace: a 512-dimensional embedding, ℓ_2 -normalized, compared by cosine similarity [6].

Several enrollment photographs become one vector, the mean of their embeddings, normalized again.

The gallery scan is linear, which is enough for a class.

A probe with no score at or above the threshold stays unrecognized.

It is not assigned the nearest identity.

The configuration threshold used in deployment is 0.68.

The clean AT&T operating point reported below is the threshold that minimized the sum of false-reject and false-accept rates on that protocol, 0.555.

The two numbers are not interchangeable.

The Eigenfaces baseline resizes a grayscale face to 64×64, equalizes the histogram, keeps 40 principal components, and classifies by Euclidean nearest neighbour [5].

It is not on the live path.

Every attempt is a row in SQLite with identity when known, time, GPS, similarity or liveness score, and status.

Status is accepted, denied at geofence, visibility, or recognition, or spoof-flagged at liveness.

The dashboard can list only the denied and spoof-flagged rows.

3.5 Interfaces

POST /checkin takes latitude, longitude, and a user id, runs (1), and returns whether the camera may open.

POST /verify takes the same location plus one image and runs the four gates.

GET /dashboard reads the log.

Enrollment is scripts/enroll.py.

There is no web enrollment form.

The browser capture page must run in a secure context, because browsers block geolocation and the camera on plain HTTP.

The public demonstration server is a CPU droplet behind nginx, served over HTTPS.

Training used a separate GPU host.

Serving does not assume a GPU at the venue.

Algorithm 1. Geofence-triggered verification

Require: fix (ϕ, λ) , venue (c, r, s) , frame source, gallery G

Require: thresholds $\tau_v = 0.70$, τ_s , attempt cap $K = 3$

1. $d \leftarrow$ Haversine distance from (ϕ, λ) to c
2. if fix missing or $d > r + s$ then
 3. return deny at geofence
4. end if
5. for $k = 1$ to K do
 6. $x \leftarrow$ capture; $v \leftarrow$ visibility(x)
 7. if $v \geq \tau_v$ then
 8. break
 9. end if
10. end for
11. if $v < \tau_v$ then
 12. return deny at visibility
13. end if
14. sample window W (about 13 frames, 1.8 s)
15. $b \leftarrow$ open–closed–open blink on EAR(W)
16. $p \leftarrow P(\text{real})$ on the highest-EAR frame
17. if not b and a recapture remains then
 18. return recapture token; do not match
19. end if
20. $\ell \leftarrow b$ or $(p \geq 0.80)$
21. if not ℓ then
 22. return spoof-flagged at liveness
23. end if
24. $e \leftarrow$ ArcFace embedding of the highest-EAR frame
25. $(u^*, \sigma) \leftarrow \text{argmax over } u \text{ in } G \text{ of } \cos(e, e_u)$
26. if $\sigma \geq \tau_s$ then
 27. return accept u^* with (d, p, σ)
28. else
 29. return deny at recognition
30. end if

4. Experiments

Three measurements were run: a clean and occluded identification comparison on the AT&T set [12], an ablation of the visibility rule on the same images, and a before-and-after check of two screen-replay photographs. A fourth check confirms that one on-site fix is accepted and one distant fix is rejected. A

four-system study that would isolate a face-only matcher, a geofence with no identity, the Jamdar pairing, and a mask-robust matcher was specified and not run. No concert, office, or multi-room trial is reported.

4.1 Protocol

The AT&T set has 40 people and 10 faces each, 400 images [12]. The recognition protocol uses 185 genuine probes and 30 impostor probes. Occlusion probes are the clean faces with a solid black overlay on the lower half, 200 images. They are not photographs of cloth masks, sunglasses, or hands. Rank-1 is the fraction of probes whose nearest gallery identity is correct. False reject and false accept use the verification threshold of that row. Liveness scores are $P(\text{real})$ on held-out photographs, not attack-presentation rates over a population.

4.2 Recognition

Table 1 compares Eigenfaces and ArcFace on the same 185 genuine probes and the same 30 impostors. The occluded rows use the same lower-half overlay and no visibility gate. Eigenfaces ranks by Euclidean distance. ArcFace ranks by cosine. The thresholds are fit separately, so they are not one shared operating point. On clean faces ArcFace is at the ceiling of this set, rank-1 of 1.000, with a false-reject rate of 0.011 (about 2 of 185) and no false accepts among the 30 impostors. Eigenfaces reaches 0.865 rank-1 and rejects 0.389 of genuine probes at its own threshold. The useful contrast is 1.000 against 0.865 on a small, aligned, studio set, not the absolute 1.000.

With the lower half covered and no visibility gate, both matchers fail in different ways. Eigenfaces rank-1 falls to 0.108 and the false-accept rate rises to 0.833. ArcFace rank-1 falls to 0.189, the genuine mean cosine falls from 0.819 to 0.071, and the false-reject rate rises to 0.984. Raw ArcFace mostly refuses the covered genuine probe. It does not become a reliable identifier of it. That is the reason the visibility gate rejects the frame instead of asking the matcher to guess.

Table 1. AT&T verification and rank-1. Occluded rows have no visibility gate. Thresholds are protocol-specific. Genuine mean cosine is given for ArcFace only.

System	Condition	Rank-1	FRR	FAR	Thr.	Cosine
Eigenfaces	clean	0.865	0.389	0.000	1824	
ArcFace	clean	1.000	0.011	0.000	0.555	0.819
Eigenfaces	occluded	0.108	0.081	0.833	3815	
ArcFace	occluded	0.189	0.984	0.000	0.511	0.071

Fig. 2 plots the four rank-1 rates in Table 1.

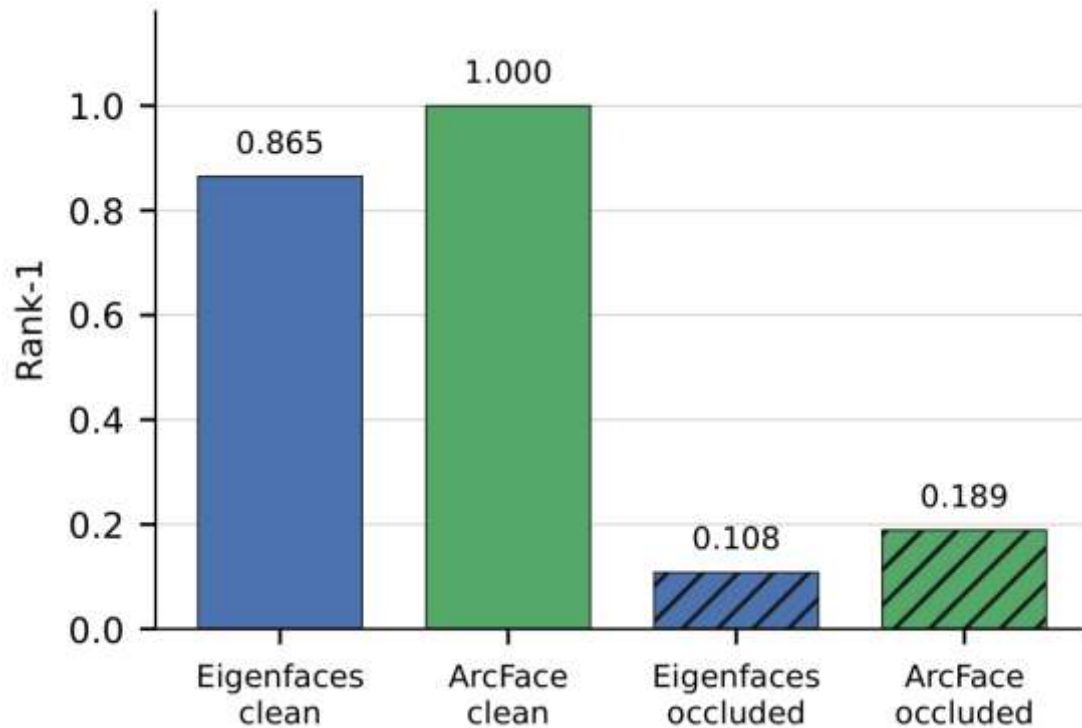


Fig. 2. Rank-1 on the AT&T protocol in Table 1. The occluded bars are the matcher with the lower face covered and no visibility gate.

4.3 Visibility Ablation

Table 2 uses the 400 clean frames and the 200 synthetic masks. The baseline composite score already catches 184 of 200 masked frames and rejects 1 of 400 clean frames. Forcing a denial on the mouth-region flag does not catch any of the remaining 16. Those 16 never raised that flag. The lower-face blackout rule catches all 16. Mean intensity and Laplacian variance on those misses were both 0, against a clean-frame fifth percentile of 94.64 and 353.40 on 399 boxed faces. Adding the Haar fallback removes the single clean miss, frame s37/4, which MTCNN did not box. The adopted gate is that combination: 200 of 200 masked frames caught, 0 of 400 clean frames rejected. The thresholds were read from this overlay distribution. They have not been re-estimated on real masks. The counts are one pass over a fixed image list. The rule is deterministic, and this paper reports no second dataset and no resampled interval.

Table 2. Visibility rules on AT&T. Catch is masked probes denied. FRR is clean frames denied.

Rule	Clean FRR	Masked catch
Baseline composite	1/400	184/200
Hard mouth flag	1/400	184/200
Lower-face blackout	1/400	200/200
Blackout + Haar fallback	0/400	200/200

Fig. 3 shows the same four rules as Table 2.

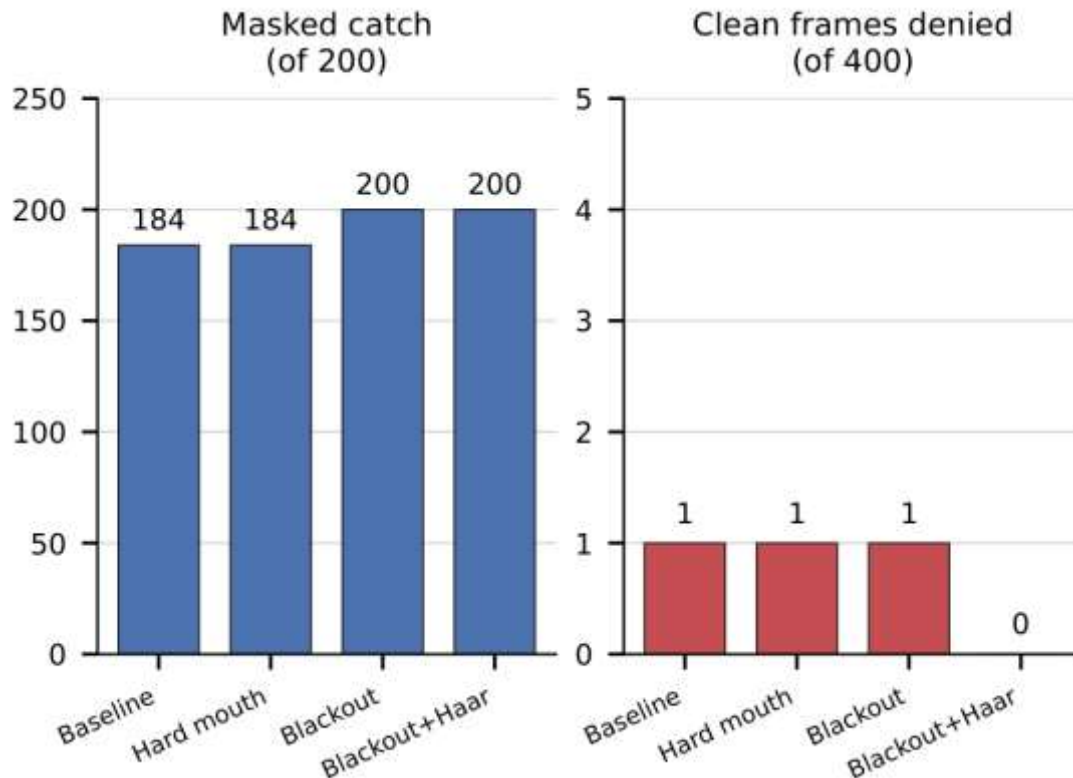


Fig. 3. Visibility-rule counts from Table 2. Masked catch is out of 200 synthetic lower-half overlays. Clean frames denied are out of 400 AT&T frames.

Table 3 restates that adopted rule as a 2-by-2 count.

Table 3. Adopted visibility rule on AT&T. The overlay row is a black rectangle on the lower half of the face. It is not a count of cloth masks, sunglasses, or hands.

Set	Denied	Accepted
Synthetic lower-half overlay	200	0
Clean AT&T frames	0	400

4.4 Liveness and the Fence

After the identity-disjoint split, NUAA training accuracy was about 99.98% and validation accuracy about 98.9%, with a best checkpoint of 99.65%. The CelebA-Spoof fine-tune reached 99.03–99.13% validation accuracy. Those rates are intra-dataset. The deployment evidence for replay is two photographs of one setup. Before the CelebA-Spoof run they scored 0.986 and 0.968 and were accepted as live. After it they scored 0.035 and 0.121 and fell below 0.80. A genuine photograph of the same capture setup scored 0.965 after the crop-margin correction. A print spoof scored 0.001. Two replay points are not an attack presentation classification error rate.

Table 4 lists the four scores.

Table 4. Four still photographs scored by the liveness model. The decision column applies to P(real) after. Empty before-cells were not reported. This table is not an attack-presentation error rate.

Case	P(real) before	P(real) after	At 0.80
Screen replay 1	0.986	0.035	reject
Screen replay 2	0.968	0.121	reject
Genuine, same setup		0.965	accept
Print		0.001	reject

Fig. 4 plots those four photographs against the live threshold of 0.80.

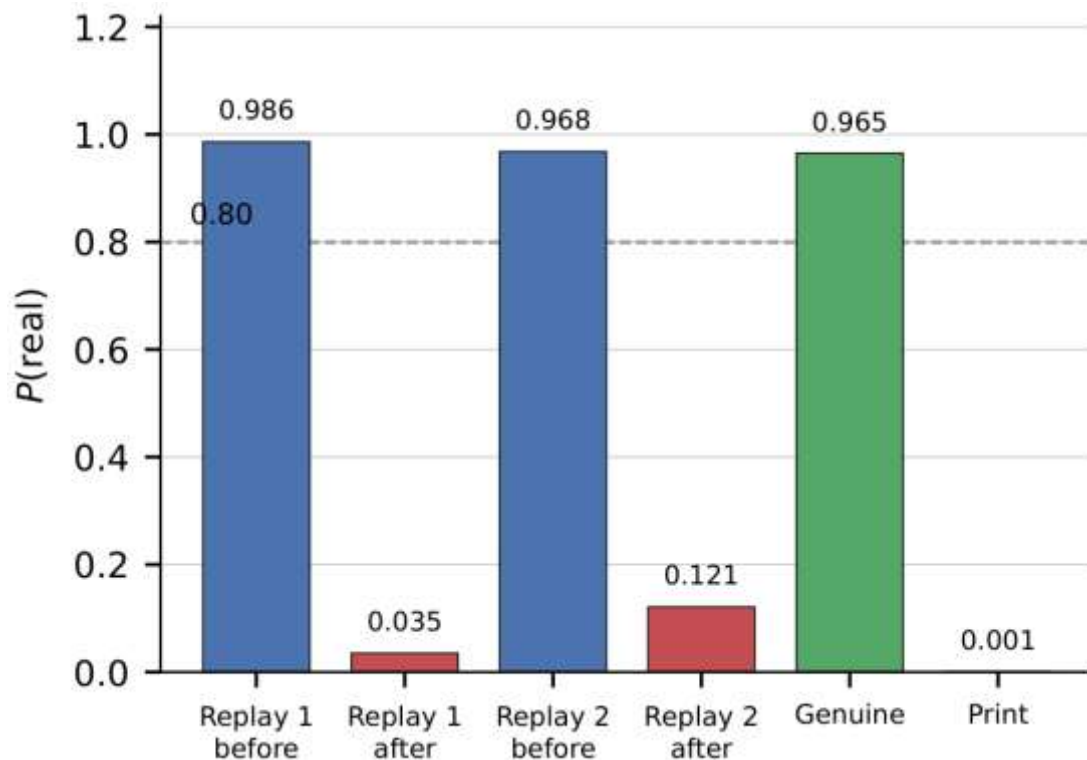


Fig. 4. Scores of four still photographs. The two screen-replay photographs are shown before and after the CelebA-Spoof fine-tune. The genuine photograph and the print have one score each. These four images are not an attack-presentation error rate.

On the demonstration server, a fix at the configured centre gives distance 0 and passes the fence. A fix 7.3 km away is denied, and the camera is not armed. That pair checks the inequality. It is not a false-accept rate outside the fence.

Table 5 lists the two fixes.

Table 5. Two demonstration fixes. The distant fix does not arm the camera. The pair is not a false-accept rate outside the fence.

Fix	Distance	Result
Configured centre	0	fence passes

Distant fix	7.3 km	denied, camera not armed
-------------	--------	--------------------------

Fig. 5 draws that pair against the 320 m cutoff.

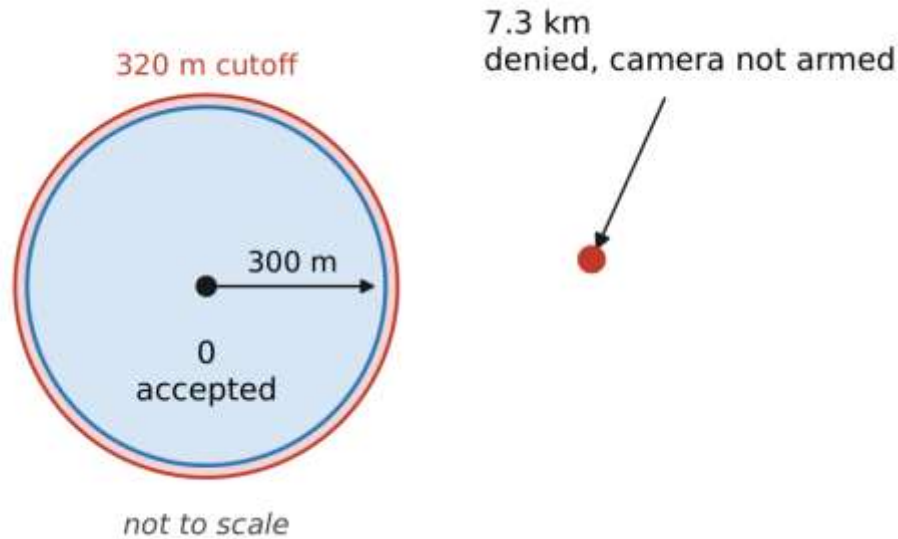


Fig. 5. Demonstration fence check. The centre fix, distance 0, passes. The fix at 7.3 km is denied and the camera is not armed. The radius is 300 m, the slack is 20 m, and the cutoff is 320 m. This pair is not a false-accept rate.

4.5 Latency

The design target is 3 s from capture to decision. Measured CPU time on the serving path is about 4–7 s. A 12.6× gap between 2.8 min on a GPU and 35.0 min on a CPU is a training-time observation for the liveness model. It is not an inference result and is not a reason to put a GPU at each venue.

5. Discussion and Limitations

The cascade does what the stage-wise numbers say. Off-site coordinates never open the camera. Synthetic lower-half coverage is stopped by a rule that looks at the lower band, after mouth-flag and detector-preprocessing variants failed. Clean-studio ArcFace outperforms Eigenfaces. Covered faces are a poor input to either matcher, which supports refusal over masked identification for this application. A reviewer can read from the log which gate fired. Jamdar et al. already pair a geofence with MobileNetV2 liveness [4]. The measurement added here is the visibility rule and the stage named in the log. It is not a liveness rate above theirs. Boe et al. and Kumar et al. report classroom appearance checks [1, 2]. The rank-1 numbers in Table 1 come from an AT&T protocol. They are not a classroom headcount.

Several limits follow directly from the protocols in this work. The occlusion result is a black rectangle placed on a 1990s studio face, so the test does not measure cloth masks, sunglasses, hands, or CCTV pose. The live demonstration already shows the mismatch: sunglasses often pass the visibility gate. A four-class visibility network was designed, but it was not trained. The liveness validation percentages sit on NUAA and on a CelebA-Spoof subset. Only two images supply the replay numbers tied to the serving crop. A real enrolled user can still receive a low live score under flat light or a static framing. That happened in demonstration use, and it is a false reject, not a spoof catch. GPS with a 300 m radius cannot decide which room a phone is in. The blink test is part of the serving liveness gate. A window is

live when a blink is detected or $P(\text{real}) \geq 0.80$. While a recapture remains, absence of a blink ends that window before recognition. The project test script uses synthetic EAR sequences for open-only, closed-only, a full transition, a one-frame dip, a broken closed run, and a window with too few valid frames. Those tests do not use a camera, and this paper reports no count of live users accepted or rejected by the blink gate. A recording that already contains a blink can still satisfy the transition. The texture scores above remain the only replay measurements, and they come from still images rather than blink windows. End-to-end latency misses the 3 s target. Quantized on-device inference and a resident model are the planned response. They were not measured for this paper.

A controlled comparison against a face-only matcher, a location-only check, a geofence-plus-liveness system without a visibility gate, and a mask-robust matcher is still required before any claim about the marginal value of each stage. Until that comparison has been run, the contribution is the implemented order, the stage-labelled log, and the measurements above.

6. Conclusion

GeoSecure-FR checks presence, visibility, liveness, and identity in that order, and it records the stage at which the attempt fails. On the evidence in hand, the fence separates an on-site fix from a fix 7.3 km away, the adopted visibility rule rejects a synthetic lower-half mask without rejecting clean AT&T frames, ArcFace is the stronger clean-set matcher, and a CelebA-Spoof fine-tune reverses two screen-replay false accepts. The limitations of this evaluation are the two fence fixes, the synthetic lower-half masks, the two replay photographs, and a CPU path that remains about 4–7 s. The same evidence does not support a presentation-attack rate, a real-mask occlusion rate, room-level location, or a latency claim below 3 s. Those are the next measurements. They are not properties of the system as evaluated.

Acknowledgement

The authors thank the Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani. This manuscript reports the GeoSecure-FR project build.

References

1. Horn B., Kok-Why N., Su-Cheng H., Palanichamy N., Elham A.A., “An Automated Face Detection and Recognition for Class Attendance”, JOIV: International Journal on Informatics Visualization, 2024, 8 (3), 1146–1153. <https://doi.org/10.62527/joiv.8.3.2967>
2. Dinesh K., Arfa T., Ajay K., Chirag S., Subramanya U.G.R., “Attendance Monitoring System Using Face Recognition with Temperature Sensor and Hand Sanitizer”, International Advanced Research Journal in Science, Engineering and Technology, 2022, 9 (7), 41–45. <https://doi.org/10.17148/IARJSET.2022.9707>
3. Zitong Y., Yunxiao Q., Xiaobai L., Chenxu Z., Zhen L., Guoying Z., “Deep Learning for Face Anti-Spoofing: A Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45 (5), 5609–5631. <https://doi.org/10.1109/TPAMI.2022.3215850>
4. Sarthak J., Akshita P., Vaishnav K., Janvi T., Ritu J., “Face Recognition-Based Smart Attendance Monitoring System Using Liveness Check and Geofencing”, Emerging Trends in Microelectronics, Communication and Intelligent Systems (ETMCIS 2024), Lecture Notes in Electrical Engineering, Springer, 2025, 1420, 287–299. https://doi.org/10.1007/978-981-96-6406-1_22

5. Matthew T., Alex P., “Eigenfaces for Recognition”, Journal of Cognitive Neuroscience, 1991, 3 (1), 71–86. <https://doi.org/10.1162/jocn.1991.3.1.71>
6. Jiankang D., Jia G., Niannan X., Stefanos Z., “ArcFace: Additive Angular Margin Loss for Deep Face Recognition”, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, 4690–4699. <https://doi.org/10.1109/CVPR.2019.00482>
7. Florian S., Dmitry K., James P., “FaceNet: A Unified Embedding for Face Recognition and Clustering”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, 815–823. <https://doi.org/10.1109/CVPR.2015.7298682>
8. Kaipeng Z., Zhanpeng Z., Zhifeng L., Yu Q., “Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks”, IEEE Signal Processing Letters, 2016, 23 (10), 1499–1503. <https://doi.org/10.1109/LSP.2016.2603342>
9. Yuanhan Z., Zhenfei Y., Yidong L., Guojun Y., Junjie Y., Jing S., Ziwei L., “CelebA-Spoof: Large-Scale Face Anti-Spoofing Dataset with Rich Annotations”, European Conference on Computer Vision (ECCV), 2020. https://doi.org/10.1007/978-3-030-58610-2_5
10. Mark S., Andrew H., Menglong Z., Andrey Z., Liang-Chieh C., “MobileNetV2: Inverted Residuals and Linear Bottlenecks”, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018, 4510–4520. <https://doi.org/10.1109/CVPR.2018.00474>
11. Tereza S., Jan C., “Real-Time Eye Blink Detection Using Facial Landmarks”, 21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia, 2016.
12. Ferdinando S., Andy H., “Parameterisation of a Stochastic Model for Human Face Identification”, 2nd IEEE Workshop on Applications of Computer Vision, 1994, 138–142. <https://doi.org/10.1109/ACV.1994.341300>