iApp Technology · Technical Whitepaper WP-2026-27

# Face Recognition: Open-Set Identification on LFW

Measured 28–29 October 2026 · iApp Technology Co., Ltd., Bangkok, Thailand

**Executive summary.** The iApp Face Recognition API names the person in a photograph from a face database the customer keeps with iApp, or answers "unknown". We enrolled one photograph of each of 1,680 people from the public LFW benchmark and searched the database with 2,581 other photographs of the enrolled people and 1,500 photographs of people who were never enrolled. At the threshold every database uses (0.50), the API named the right person for 84.11% of the enrolled people's photographs (95% CI [82.72, 85.59]), named a wrong enrolled person for 0.31% (8 searches), and answered "unknown" for the rest. It named someone for 0.13% of the strangers (2 of 1,500). Ignoring the threshold, the best match was the right person for 97.60% of searches. These figures are for one enrolment photograph per person, the hardest setting; the documentation recommends three to five. Removing a person through the API deleted their face data from the service's storage.

## 1 Introduction

Face identification (1:N) compares one face with every face in a database and returns the best match, or "unknown" when no enrolled face is similar enough. Applications include access control, attendance and visitor registration. Two errors matter: naming the wrong enrolled person, and naming someone who is not enrolled at all. Both depend on the recognition threshold of the database, a score from 0 to 1 that the best match must exceed.

This paper measures both errors, and the rate of correct identification, on a public benchmark at the threshold every database uses (0.50), over the whole threshold range, and at the stricter values an application can apply to the returned score. It also checks, on the live service, that removing a person deletes their face data.

## 2 Test Data and Protocol

LFW [1] holds 13,233 news photographs of 5,749 people. We used it as follows:

- **Database.** Each of the 1,680 people with at least two photographs is enrolled with one photograph, the first in the published order. 0 enrolment photographs had no face the service accepted.
- **Genuine searches.** Up to two other photographs of each enrolled person, chosen with seed 1337: 2,581 searches.
- **Stranger searches.** 1,500 photographs, chosen with seed 1337, of people who have only one photograph in LFW and are therefore never enrolled.

The photographs are of public figures, mostly adults of European descent. One enrolment photograph per person is the hardest setting; the documentation recommends three to five. None of the images is shown in this paper, and the test database contains no names: people were enrolled under pseudonyms.

## 3 Method

**Full-scale measurement.** The recognition service computes a face representation for every image and compares it with each enrolled face. We sent every image of the protocol once to the production service

that the public `/v3/store/ekyc/face-recognition` endpoints call, one request at a time, obtained the representation the service computes, and applied the service's own decision rule to the 1,680-person database: the enrolled face with the highest score is returned when its score exceeds the threshold, otherwise "unknown". We did not enrol 1,680 people in the live service, to keep the load on a service that customers use to a minimum.

**Check against the live API.** To confirm that this reproduces what the API returns, we created one dedicated test database through `https://api.iapp.co.th`, enrolled 20 people (20 accepted), and sent 53 searches to the production service and the same 53 searches through the public API. The service returned the predicted name for 53 of 53 searches, and through the public API for 53 of 53; the largest difference between a returned score and the predicted score was 0.0000.

**Removal.** Every enrolled person was then removed through the public API (20 of 20 removals succeeded), after which the API reported that the test database no longer exists (HTTP 421 to both the count and the export request). We also read the service's storage directly, with operator access, at three points (Table 3). After the removals no stored face record of the test database remained. Its password, stored as a salted hash, remains until the database itself is deleted; we then deleted it.

**Metrics.** For genuine searches: the share for which the API returns the right person (the identification rate), a wrong enrolled person, or "unknown". For stranger searches: the share for which the API returns any name (the false-positive identification rate). Searches in which the service finds no face count as "unknown". Rank-1 accuracy ignores the threshold. 95% confidence intervals are percentile bootstrap intervals (2,000 resamples, seed 1337).

## 4 Results

**Table 1:** Identification when a name is accepted only above each threshold, 1,680-person database, 2,581 genuine and 1,500 stranger searches. The first row is the database threshold; the others are stricter values an application can apply to the returned score.

| Threshold | Right person (%) | Wrong person (%) | Unknown (%) | Stranger named (%) |
|---|---|---|---|---|
| 0.50 (default) | 84.11 | 0.31 | 15.58 | 0.13 |
| 0.55 | 74.20 | 0.27 | 25.53 | 0.13 |
| 0.60 | 61.29 | 0.15 | 38.55 | 0.13 |
| 0.65 | 45.95 | 0.04 | 54.01 | 0.13 |
| 0.70 | 29.06 | 0.04 | 70.90 | 0.07 |
| 0.85 | 2.25 | 0.04 | 97.71 | 0.07 |

At the database threshold the 95% confidence intervals are [82.72, 85.59] for the right person and [0.00, 0.33] for strangers named.

**Table 2:** Threshold-free measures and fixed false-positive rates.

| Measure | Value |
|---|---|
| Rank-1: best match is the right person (2,581 searches) | 97.60% [96.94, 98.18] |
| Right person among the five best matches | 98.18% |
| Right person when 1% of strangers are named (threshold 0.40) | 93.76% |
| Right person when 0.1% of strangers are named (threshold 0.67) | 40.57% |
| Searches with no face found: genuine / stranger | 0 / 0 |

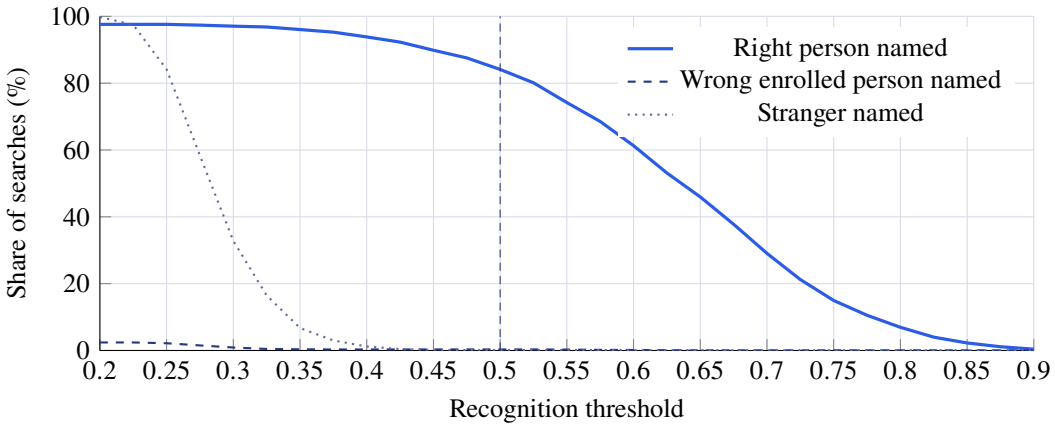



**Figure 1:** The three outcomes against the threshold; the dashed vertical line marks the database threshold 0.50.

## 5 Discussion

**At the default threshold the API rarely names the wrong person.** With 1,680 people enrolled, a search named a wrong enrolled person 8 times in 2,581 (0.31%) and named a stranger 2 times in 1,500 (0.13%). The cost of that caution is that 15.58% of genuine searches return "unknown": with a single, uncontrolled enrolment photograph, a news photograph taken years apart or from another angle often scores below 0.50. An application should let the user retry, and enrol several photographs per person.

**Ranking is better than the decision.** The right person was the best match for 97.60% of searches and among the five best for 98.18%. Lowering the threshold to 0.40 would find the right person for 93.76% of searches, at the price of naming 1% of strangers; raising it to 0.67 would cut that to 0.1% and find 40.57%. The default of 0.50 sits between the two (Figure 1).

**A stricter threshold mostly adds "unknown" answers.** An application can accept a name only above a higher score than 0.50. Above 0.50 the API already names a wrong person or a stranger rarely, so a stricter value mostly turns right answers into "unknown": at 0.85, only 2.25% of genuine searches named the right person and 97.71% returned "unknown". Recognition scores of the same person in two photographs rarely exceed that value.

**The claim the evidence supports:** on LFW, with one enrolment photograph per person and 1,680 people enrolled, the API identifies the right person in 84.11% of searches at its default threshold, names a wrong enrolled person in 0.31% and a stranger in 0.13%.

## 6 Removal and Stored Data

**Table 3:** The service's storage for the test database, read on the service with operator access.

| | After enrolment | After removal | After deletion |
|---|---|---|---|
| Stored face records (backup copy) | 20 | 0 | 0 |
| Database file | yes | no | no |
| Database entry (password, thresholds) | yes | yes | no |
| Passwords stored in plain text, all databases | 0 | 0 | 0 |

The API stores, for each enrolled face, the name, an identifier and the face representation; images are not kept. Removing a name deletes its face representations from the service at once, including the service's

own backup copy. A database's password is stored only as a salted hash; no database's password was stored in plain text (Table 3). Disaster-recovery backups that may still hold removed data are replaced within seven days.

## 7 Production Performance

Through `https://api.iapp.co.th`, from a client in Bangkok, a search in the 20-person test database took a median of 0.59 s end to end (90th percentile 0.89 s). Search time grows with the size of the database, as every enrolled face is compared. A search costs 0.3 IC and an enrolment 0.1 IC.

## 8 Limitations

- **Population and images.** LFW is news photographs of public figures, mostly adults of European descent; it does not measure accuracy on Thai faces or on access-control cameras.
- **One enrolment photograph.** With three to five photographs per person, as the documentation recommends, results would be expected to improve; this was not measured.
- **Database size.** Error rates grow with the number of enrolled people. These results hold for 1,680 people.
- **Method.** The full-scale figures apply the service's decision rule to the representations the service returned, not 1,680 live enrolments; the live check covered 20 enrolled people and 53 searches. The storage check reads the service's files with operator access, which a customer cannot repeat.
- **Public benchmark.** LFW has been public since 2007; we cannot rule out that its images were seen when the recognition model was built.

## 9 Reproducibility

LFW is public. The scripts `bench.py` and `gen_results.py`, the per-search results (identities as indices; no images, no face data) and the seeds are available from iApp Technology on request: sale@iapp.co.th.

## References

[1] G. B. Huang, M. Ramesh, T. Berg, E. Learned-Miller. *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments*. University of Massachusetts, Amherst, Technical Report 07-49, 2007.

**About iApp Technology.** iApp Technology Co., Ltd. is a Thai artificial-intelligence company providing OCR, document understanding, speech, identity verification and large-language-model services for the Thai language.
Web: iapp.co.th · E-mail: sale@iapp.co.th · Tel: +66 2 124 4041