Morphing Attack Potential

Matteo Ferrara

Department of Computer Science and Engineering

University of Bologna - Italy

iMARS

image manipulation attack
resolving solutions

International Face Performance Conference
15-17 November 2022

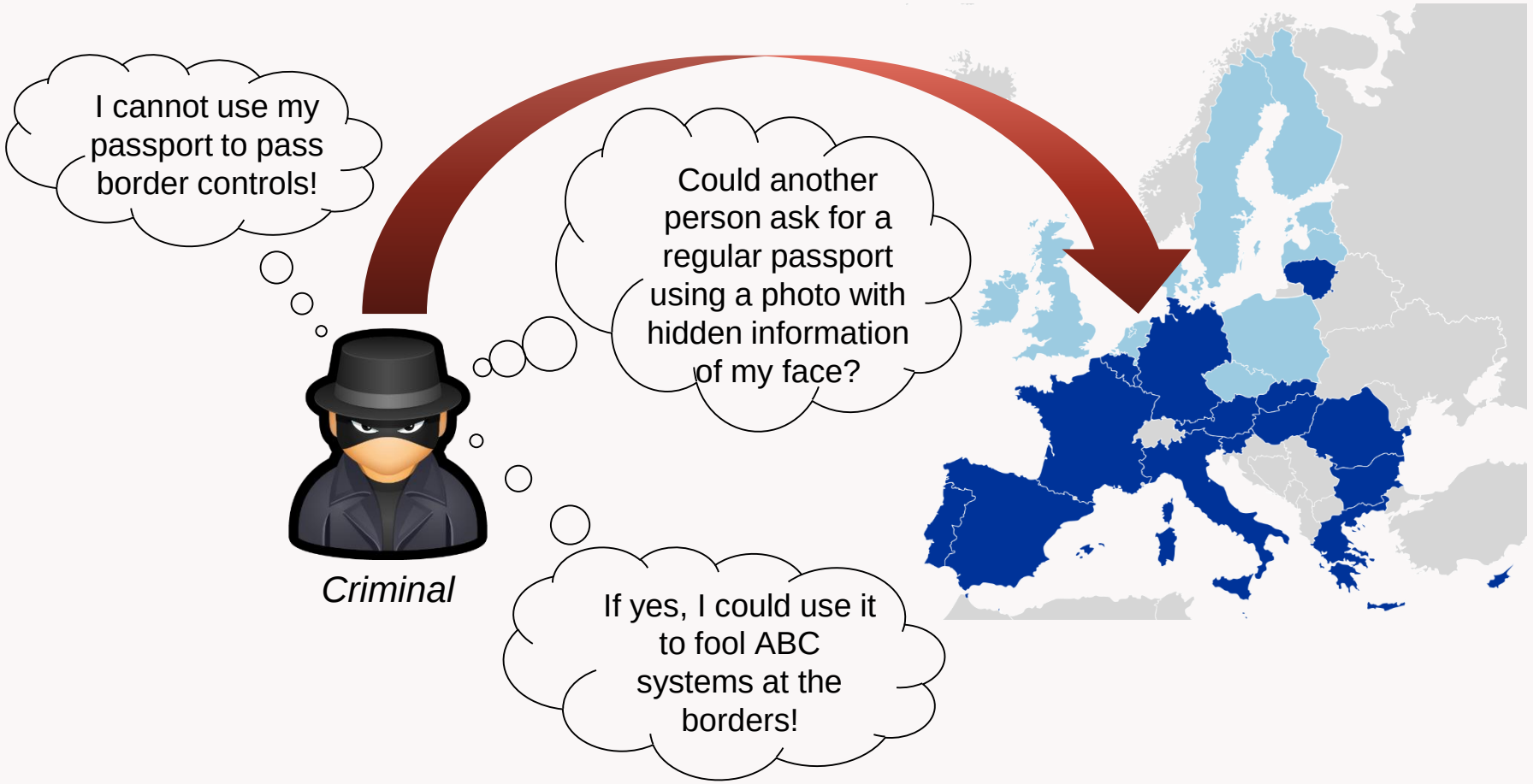
What is morphing?

“In computer graphics and animations, morphing is a special effect that transforms an image into another through a seamless transition”

https://noahmjacobs.com/computer-vision/face-morphing/

The iMARS project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 883356

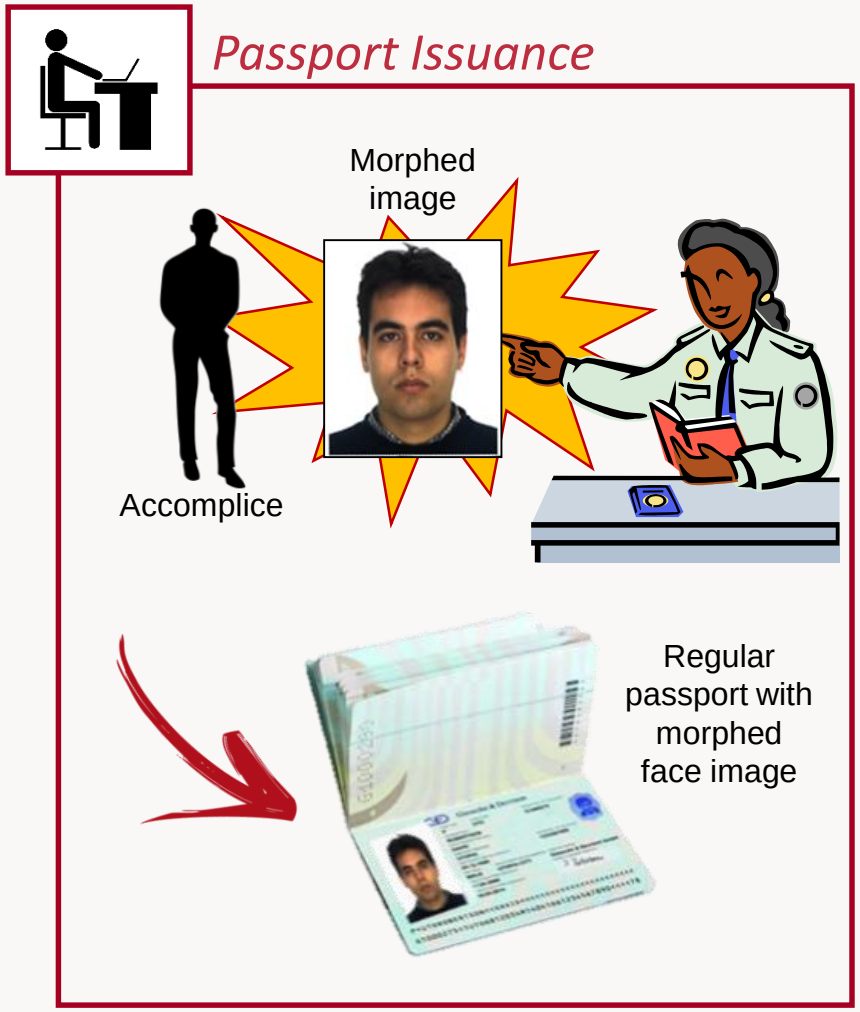
The morphing attack

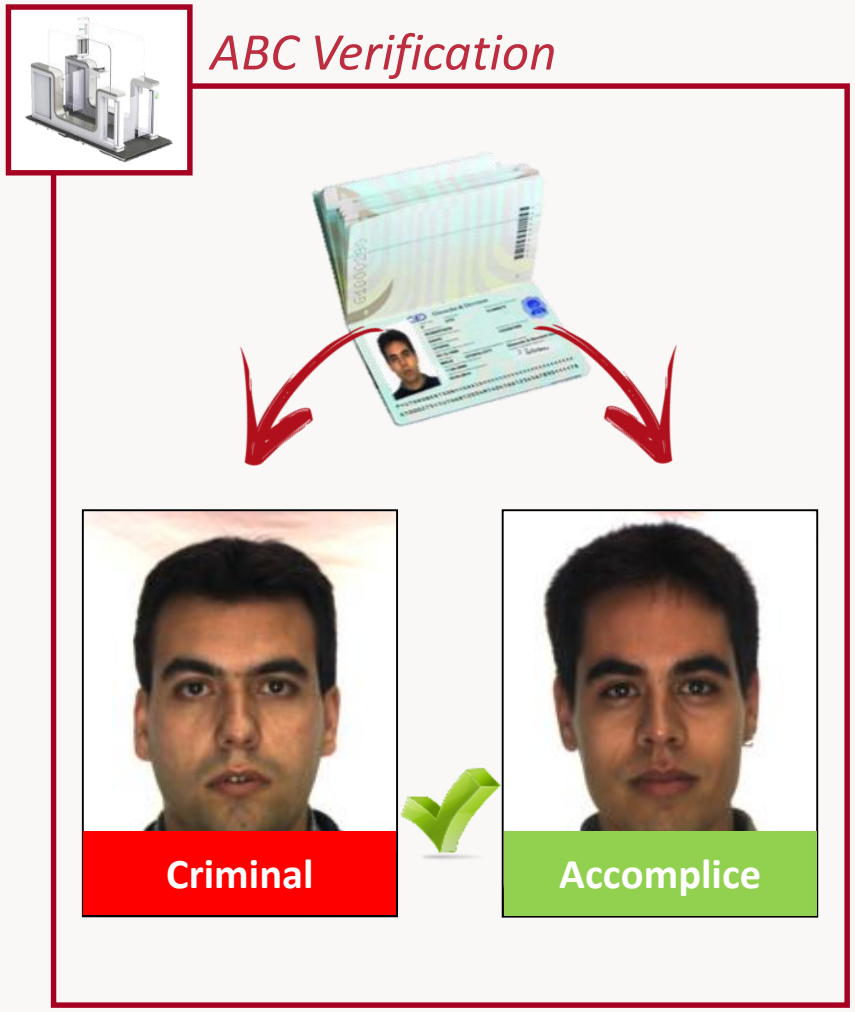


The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

The morphing attack (2)

If a double-identity face image can be enrolled in the chip, two subjects can share the same document





The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

The morphing attack (3)

- The issued document is perfectly regular.
- The attack does not consist of altering the document content but in deceiving the officer during document issuing. For this reason the morphed photo ID must be very similar to the applicant.
- The document released will thus pass all the integrity checks performed at the gates.
- It has been proved that:
  1. it is possible to create a realistic morphed image;
  2. the morphed image is able to deceive the officer;
  3. state-of-the-art Face Recognition Systems (FRSs) can be easily fooled.

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

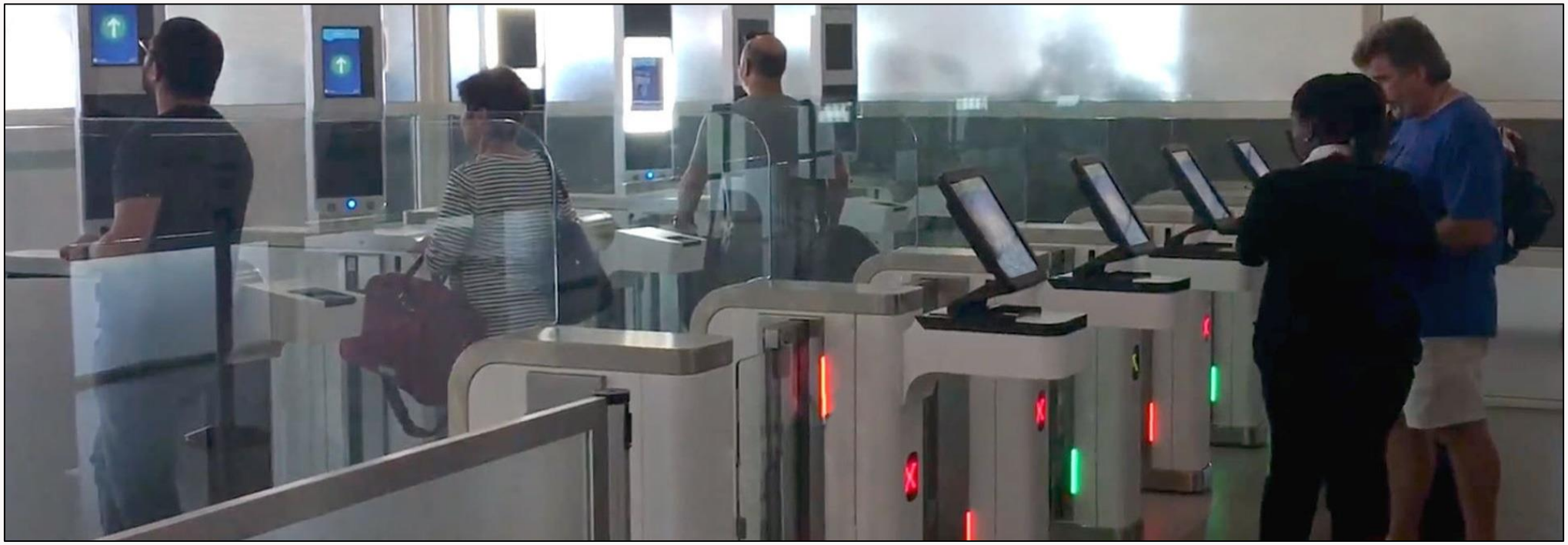
A real case

On October 2018, German activists used a morphed image of Federica Mogherini (High Representative of the European Union for Foreign Affairs and Security Policy) and a member of their group to get a genuine German passport.

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

ABC gate scenario

- The verification process at an ABC gate is performed by comparing the document image against multiple consecutive frames acquired live.



The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

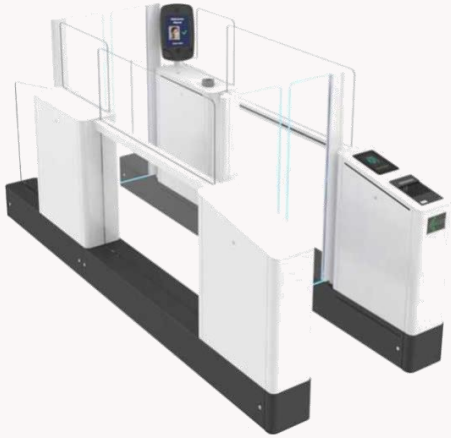
ABC gate scenario (2)

- ABC gates of different manufacturers use different FRSs.
- Different FRSs use a different number of live frames during the verification.

Airport A

- FRS 1
- 5 live frames acquired
- Verification is based on the most similar live frame

Airport B



- FRS 2
- 1 live frame acquired
- Verification is based on the single similarity score

Airport C

- FRS 3
- 10 live frames acquired
- Verification is based on the 5 highest quality frames

Airport D

- FRS 4
- 15 live frames acquired
- Verification is based on the similarity of all frames

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

How measure the vulnerability to morphing of a FRS?

- When is a morphing attack considered successful?
  - Only if all contributing subjects are successfully matched against the morphed sample.
- The vulnerability to morphing is usually measured on specific databases of morphed images.
- It is quantified as the proportion of morphed images that are erroneously verified as bona fide with all contributing subjects.
- Two metrics have been introduced for vulnerability assessment.

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

Mated Morph Presentation Match Rate (MMPMR)

- A morphing attack succeeds if the morphed image can be successfully verified against at least one of the probe images of each subject.

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

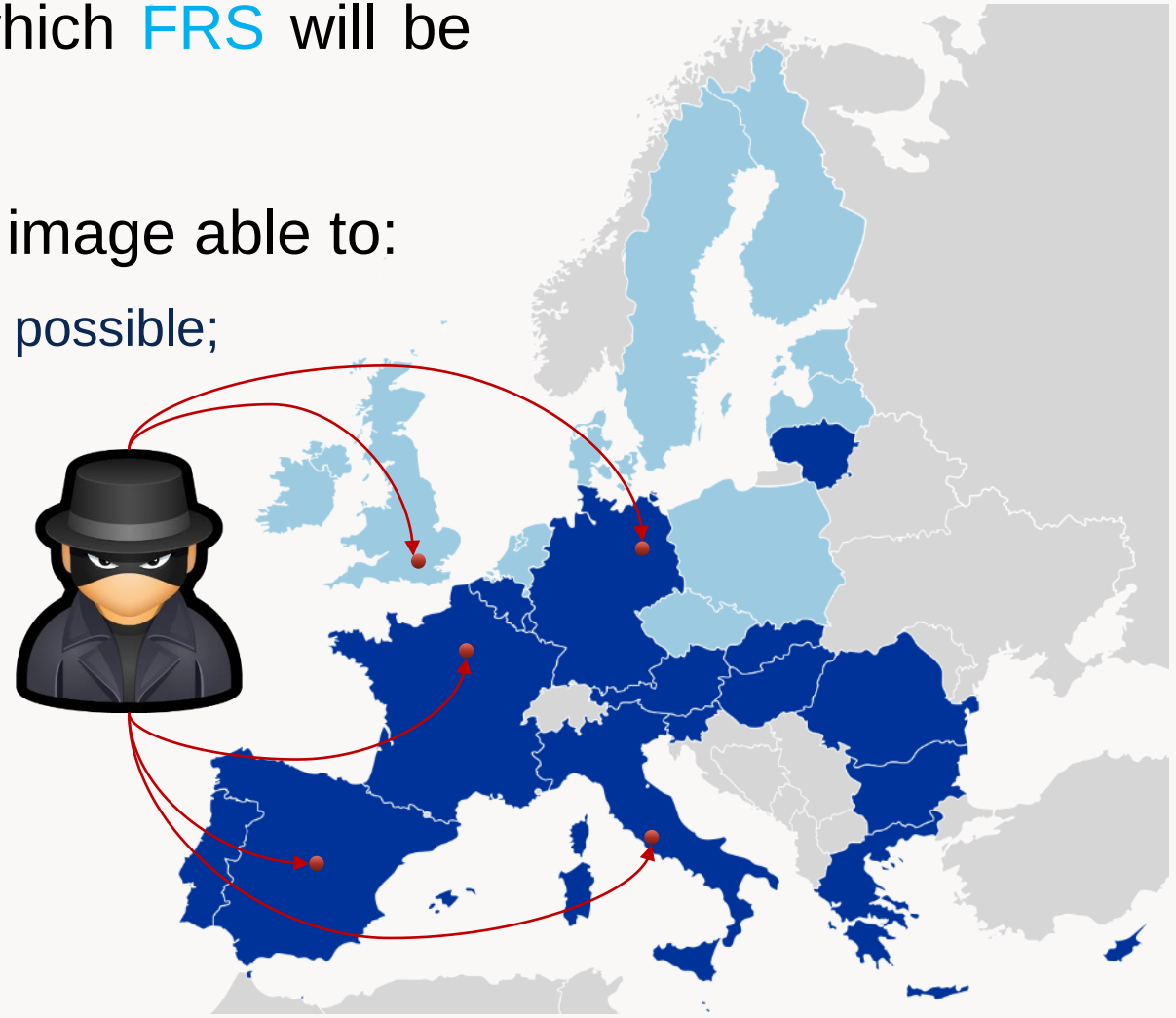
Fully Mated Morph Presentation Match Rate (FMMPMR)

- A morphing attack succeeds if the morphed image can be successfully verified against all probe images of each subject.

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

The criminal's perspective

- A criminal does not know in advance which FRS will be used at the destination airport.
- The criminal goal is to create a morphed image able to:
  - be verified against as many probe images as possible;
  - fool as many FRSs as possible.
- How to evaluate the comprehensive attack potential of:
  - a morphed image?
  - a morphing process?
  - a dataset of morphed images?



The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

Limits of available vulnerability metrics

- The MMPMR and FMMPMR can only partially estimate the attack potential.
- They do not take into account:
  - multiple FRSs (*generality*);
  - a variable number of verified probe images (*robustness*).
- To extend these concepts we proposed a new metric called Morphing Attack Potential (MAP) that considers a variable number of attempts (frames acquired live at the gate) and multiple FRSs.

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

Morphing attack potential

- Given a dataset of morphed images $\mathbb{M}$, m probe images for each contributing subject and n FRSs to evaluate, MAP is defined as a matrix of size $m \times n$ whose element $\mathrm{MAP}[r, c]$ reports the proportion of morphed images successfully verified with both contributing subjects with at least r probe images by at least c FRSs.

85% of the morphed images fool at least one FRS with at least one of the probe images.

Example:
$m = 5$
$n = 4$

generality

robustness

MAP		# FRSs (*c*)			
		1	2	3	4
# attempts (*r*)	1	85%	73%	60%	48%
	2	80%	68%	55%	43%
	3	75%	63%	50%	38%
	4	70%	58%	45%	33%
	5	65%	53%	40%	28%

28% of the morphed images fool all four FRSs with all five probe images (MOST DANGEROUS).

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

Experimental results

- SOTAMD database:
  - 150 different subjects.
  - 10 probe images for each subject (m).
  - 7 different morphing algorithms.
  - Digital and printed & scanned formats.
  - 5748 high quality morphed images.
- 4 pre-trained models as FRSs (n):
  - ArcFace
  - Dlib
  - Facenet
  - VGG-Face
- Testing protocol:
  - Each morphed image has been compared against all 10 probe images of both subjects using the 4 FRSs.
  - The threshold of each FRS has been fixed to ensure a $FMR = 0.1\%$.

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

Experimental results (2)

MAP computed on the entire SOTAMD dataset (5748 morphed images)

# Attempts (r)	# FRSs (c) = 1	2	3	4
1	39.6%	16.8%	6.1%	1.3%
2	32.9%	12.6%	4.5%	0.9%
3	29.3%	10.5%	3.5%	0.5%
4	26.0%	8.4%	2.4%	0.4%
5	23.4%	6.8%	1.9%	0.2%
6	19.7%	5.6%	1.4%	0.1%
7	16.4%	4.6%	1.0%	0.1%
8	13.7%	3.7%	0.7%	0.1%
9	11.5%	2.6%	0.3%	0.0%
10	7.6%	1.6%	0.0%	0.0%

FNMR measured on the SOTAMD bona fide images

FRS	FNMR@FMR=0.1%
ArcFace	6.0%
Dlib	29.8%
Facenet	27.5%
VGG-Face	31.4%

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

Experimental results (3)

MAP computed on the morphed images obtained with 2 of the most promising morphing algorithms used in SOTAMD (700 and 1359 morphed images)

# Attempts (r)	# FRSs (c) = 1	2	3	4
1	51.6%	19.1%	5.4%	0.9%
2	43.4%	13.7%	3.0%	0.3%
3	37.1%	10.0%	1.7%	0.1%
4	31.6%	8.1%	0.7%	0.0%
5	28.3%	6.9%	0.6%	0.0%
6	21.4%	5.6%	0.3%	0.0%
7	16.6%	4.1%	0.1%	0.0%
8	11.7%	3.4%	0.0%	0.0%
9	8.7%	2.1%	0.0%	0.0%
10	5.0%	1.0%	0.0%	0.0%

# Attempts (r)	# FRSs (c) = 1	2	3	4
1	38.2%	16.5%	6.3%	1.9%
2	32.3%	12.2%	4.7%	1.3%
3	28.7%	9.9%	3.8%	1.0%
4	24.9%	7.4%	2.6%	0.9%
5	22.4%	5.8%	2.1%	0.4%
6	18.8%	4.9%	1.6%	0.4%
7	15.6%	4.1%	1.0%	0.1%
8	13.2%	3.5%	0.9%	0.1%
9	10.7%	2.6%	0.2%	0.0%
10	7.6%	1.8%	0.1%	0.0%

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

MAP usefulness

- MAP can also be useful to answer the following questions:
  - What is the impact of one morphing method on a set of FRSs?
    - using a dataset containing morphed images generated by such algorithm.
  - What is the vulnerability of one (operational) FRS to morphing?
    - computing a MAP-matrix with a single column.
  - What is the impact of a specific factor (e.g., morphing factor, subject age, JPEG compression, print and scan process, etc.) on the attack potential of morphing?
    - using a dataset containing only morphed images with the specific factor.

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

MAP as a new ISO standard

- We are working on a New Work Item Proposal to submit to ISO/IEC JTC1 SC37
  - **Title:** Vulnerability of biometric recognition systems with respect to morphing attacks.
  - **Scope:** This standard establishes requirements for biometric recognition systems that could become subject to morphing attacks.

The iMARS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356

Contacts & references

Matteo Ferrara - matteo.ferrara@unibo.it

Biometric System Laboratory - biolab.csr.unibo.it

iMARS Project - imars-project.eu

M. Ferrara, A. Franco, D. Maltoni and C. Busch, "Morphing Attack Potential", in proceedings *IEEE International Workshop on Biometrics and Forensics (IWBF)*, Salzburg, Austria, April 2022.

iMARS
image manipulation attack
resolving solutions

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883356.
This presentation reflects only the author's views, and the Commission is not liable for any use that may be made of the information contained therein.