IEEE ICASSP 2022, Paper ID: 9141

Are GAN-based Morphs Threatening Face Recognition?

Eklavya Sarkar[1,2], Pavel Korshunov[1], Laurent Colbois[1,3], and Sébastien Marcel[1,3]

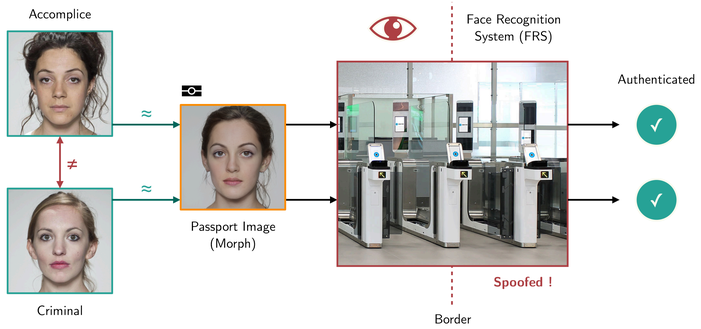
[1]Idiap Research Institute, Martigny, Switzerland [2]École polytechnique fédérale de Lausanne, Switzerland [3]University of Lausanne, Switzerland

Aims

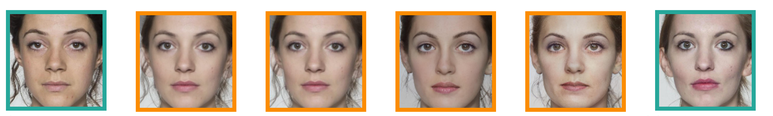
- Assess the level of vulnerability of four existing SOTA face recognition (FR) systems against four different types of morphing attacks.

Morphing Attacks

- When two individuals' face images are combined into a single 'morphed' image using a morphing algorithm.
- A threat to any biometric FR system where reference in an identity document can be altered.



Morph Generation

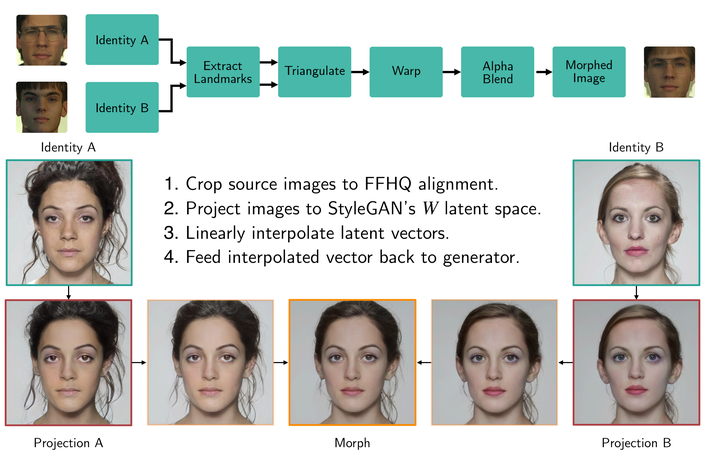


Landmark based morphs:

- b) OpenCV
- c) FaceMorpher

GAN based morphs:

- d) StyleGAN2
- e) MIPGAN-II



Evaluation Protocols

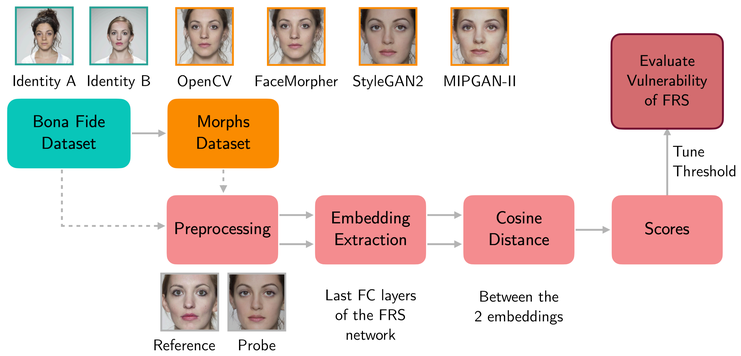
Databases:

- *FERET*: standard dataset commonly used in papers on morphing attack detection with a large number of images of different identities.
- *FRLL*: close-up frontal face images of 1350×1350 resolution, shot under *uniform* illumination with large varieties in ethnicity, pose, and expression.

Face Recognition Systems (accuracy on LFW dataset):

- *FaceNet* (99.6%)
- *ArcFace*: (99.5%)
- *VGG-Face*: (98.5%)
- *Inter-Session Variability (ISV)*: trained on MOBIO dataset.

Pipeline:



Verification categories:

- *Genuine user*: probe and claimed identity both correctly belong to the user.
- *Zero-effort impostor*: probe belongs to the user, but the claimed identity corresponds to a different enrolled user.
- *Morph attack impostor*: probe matches the claimed identity but does not correspond to the user.

Metrics:

- *False Match Rate (FMR)*: proportion of zero-effort impostors that are falsely authenticated.
- *False Non-Match Rate (FNMR)*: proportion of genuine users which are falsely rejected.
- *Mated Morph Presentation Match Rate (MMPMR)*: proportion of morphs attacks impostors accepted by the face recognition system.

Scenarios:

- Bona Fide (BF): both reference and probe images are genuine.
- Morphing Attack (MA): morphs are introduced to the FR system with an intention of spoofing.
  - *Morphs as references*: FR system is hijacked during enrollment process.
  - *Morphs as probes*: similar to presentation attack scenario.

Experimental Results

Table 1. MMPMR @ FMR = 0.1% (Morphs as references — Morphs as probes) [%]

Tools	FRS	FRLL	FERET
OpenCV	FaceNet	83.3 — 72.0	41.1 — 40.6
	Arcface	59.8 — 73.8	34.6 — 35.2
	VGG	39.7 — 48.6	22.0 — 21.0
	ISV	59.8 — 97.8	44.8 — 58.4
FaceMorpher	FaceNet	64.5 — 68.2	39.9 — 40.3
	Arcface	57.6 — 75.3	34.1 — 34.8
	VGG	23.4 — 47.1	20.5 — 18.3
	ISV	56.1 — 96.1	42.6 — 56.5
StyleGAN2	FaceNet	5.9 — 11.0	1.6 — 1.3
	Arcface	9.8 — 18.3	2.4 — 2.5
	VGG	3.0 — 9.1	2.0 — 1.5
	ISV	9.2 — 43.6	2.7 — 3.4
MIPGAN-II	FaceNet	47.2 — 62.7	32.9 — 32.3
	Arcface	32.0 — 46.5	26.0 — 25.1
	VGG	15.9 — 30.4	14.5 — 13.2
	ISV	3.6 — 23.7	7.3 — 9.6

- StyleGAN2-morphs do *not* pose a significant threat to SOTA face recognition systems, compared to landmark-based morphs, despite being of higher visual quality, and with very few ghosting artefacts.
- The more accurate face recognition system is the more vulnerable it is to morphing attacks. See: FaceNet vs VGG.
- The quality of original images used to create morphs may lead to more threatening morphs in the presentation attack scenarios, rather than when attacking FR systems from the inside.

Conclusion

- 'Classical' morphs are much more threatening to automated FR systems than GAN-based morphs.
- FR systems which are better at recognition are also more vulnerable to morphing attacks.

Release

We provide:

- An open-source morphing tool for generating the morphing attacks.
- An open source package for running the evaluation experiments.
- The generated and used datasets of morphed images.