Vulnerability Analysis of Face Morphing Attacks from Landmarks and Generative Adversarial Networks

—

Eklavya SARKAR
Research Intern, Biometrics Security and Privacy, Idiap Research Institute

Content

Content

- My journey to Idiap

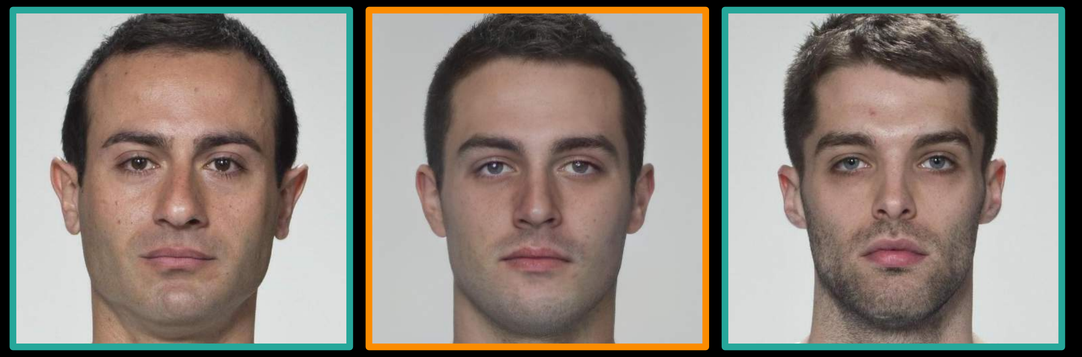
Content

- My journey to Idiap

- Problem
- Morph Generation
- Evaluation Protocol
- Experimental Results
- Conclusion

Morphing Attacks

A Threat to Biometric Systems

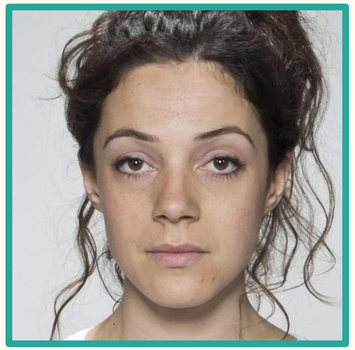


Karras, T. *et al.*, 2020. Analyzing and improving the image quality of stylegan. CVPR.

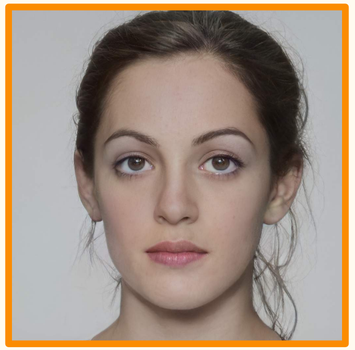
Problem

Problem

Morphing Attack: When two individuals' face images is combined into a single 'morphed' image using a morphing algorithm.



Identity A



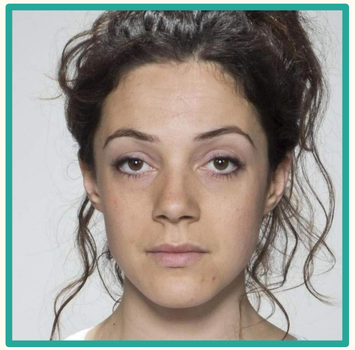
Morph

Identity B

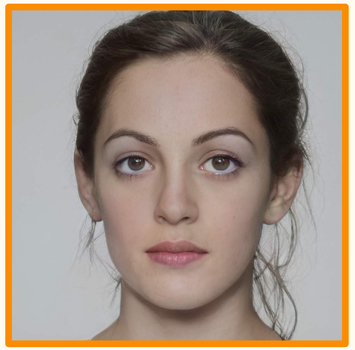
Problem

Morphing Attack: When two individuals' face images is combined into a single 'morphed' image using a morphing algorithm.

- A threat to any biometric system where reference in an identity document can be altered.



Identity A



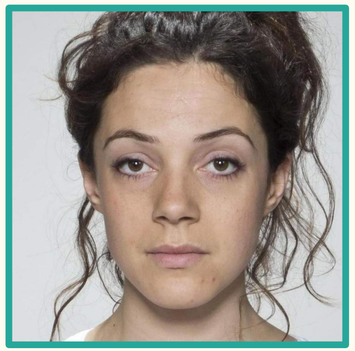
Morph

Identity B

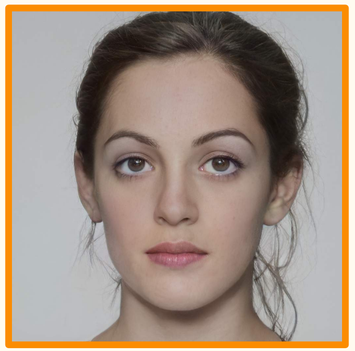
Problem

Morphing Attack: When two individuals' face images is combined into a single 'morphed' image using a morphing algorithm.

- A threat to any biometric system where reference in an identity document can be altered.
- Presents an important issue in systems relying on identity documents.



Identity A



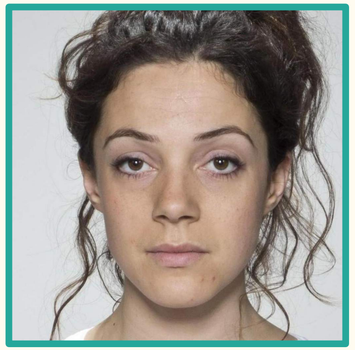
Morph

Identity B

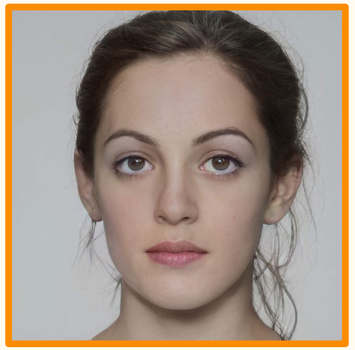
Problem

Morphing Attack: When two individuals' face images is combined into a single 'morphed' image using a morphing algorithm.

- A threat to any biometric system where reference in an identity document can be altered.
- Presents an important issue in systems relying on identity documents.
  - Automatic border control



Identity A



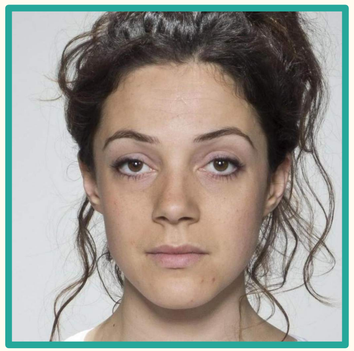
Morph

Identity B

Morphing Attack - Automatic Border Control

Morphing Attack - Automatic Border Control

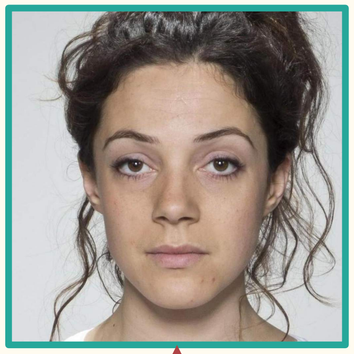
Accomplice



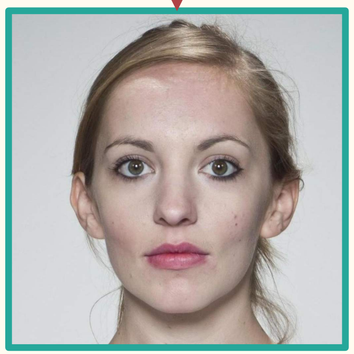
Criminal

Morphing Attack - Automatic Border Control

Accomplice

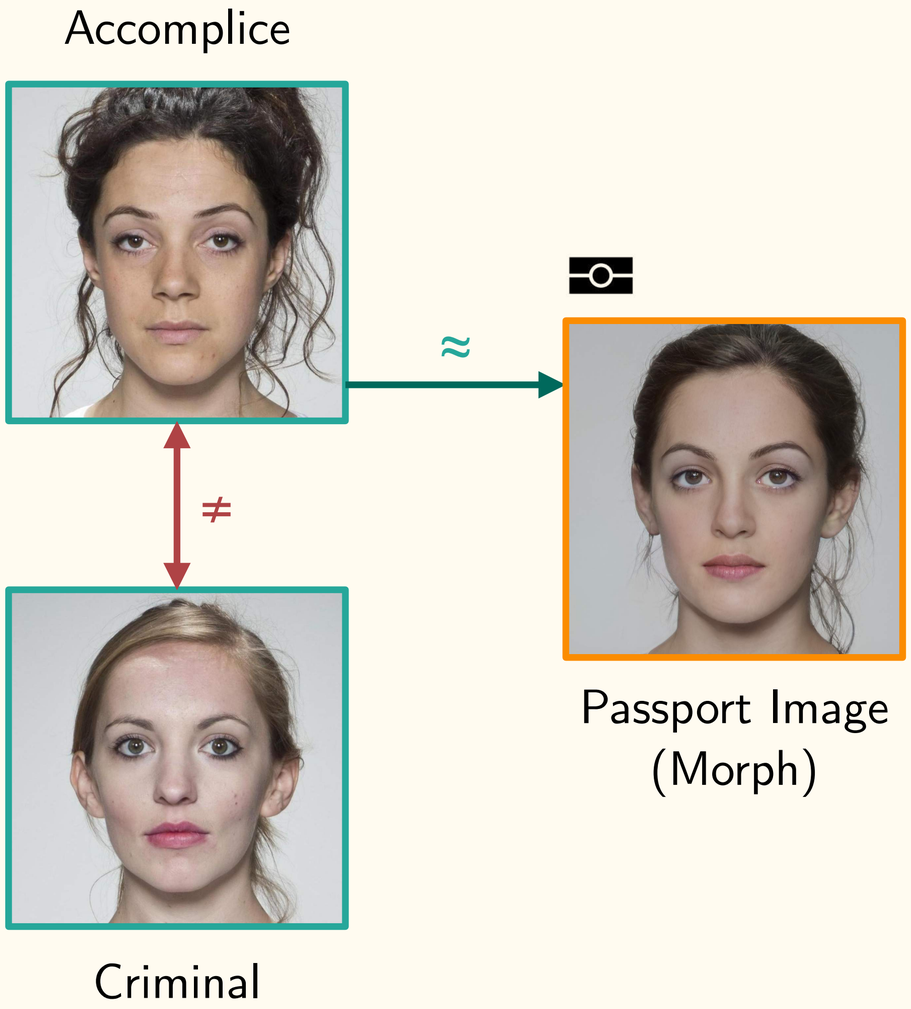


≠

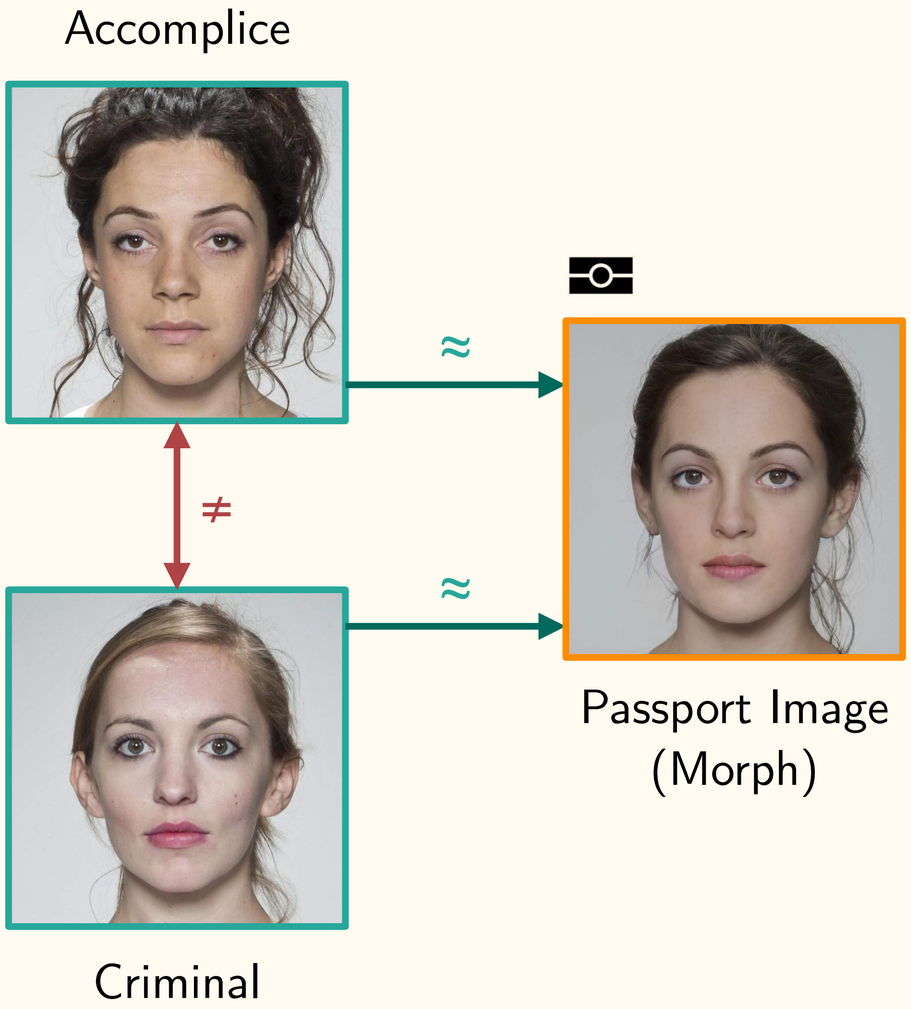


Criminal

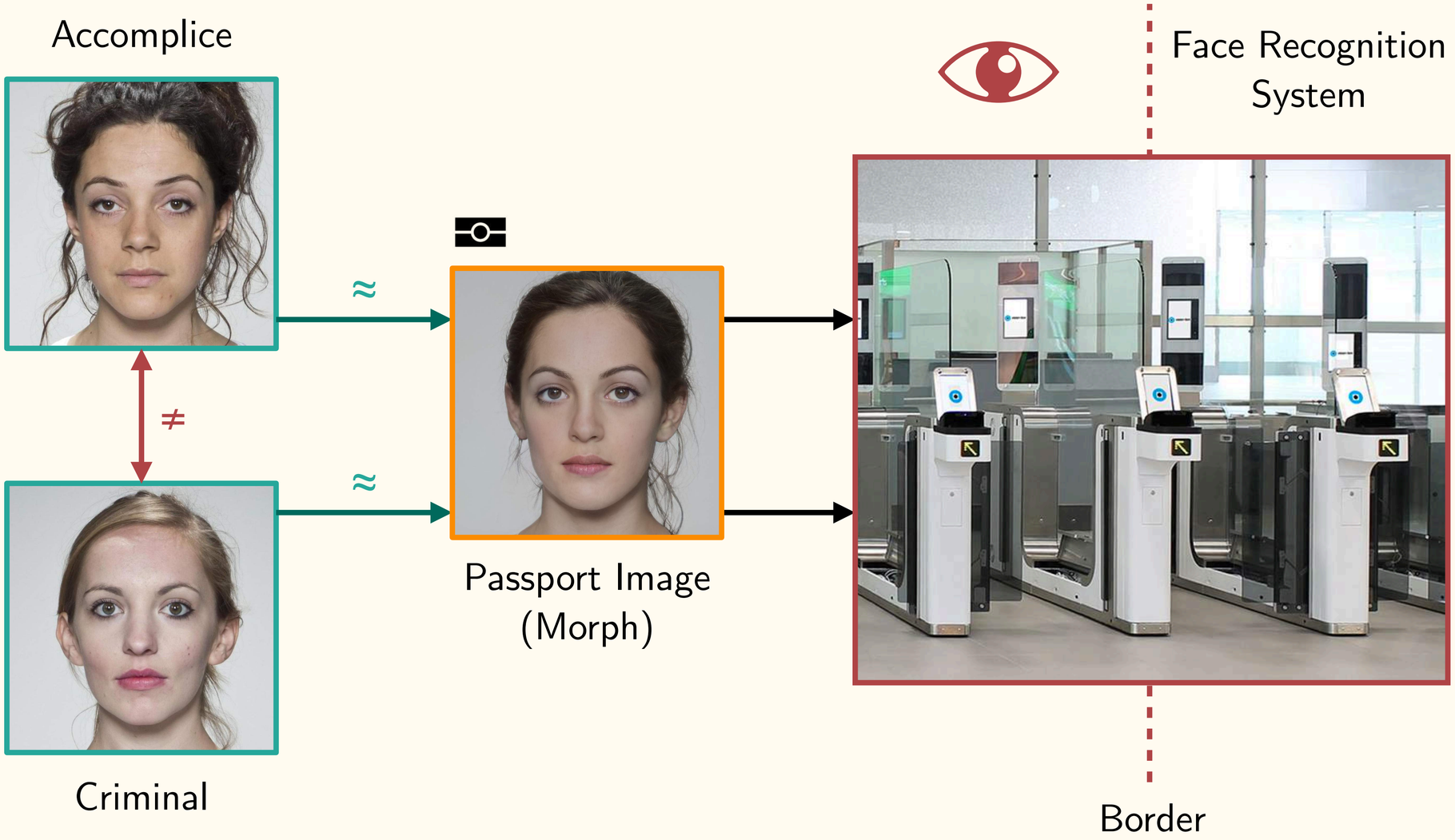
Morphing Attack - Automatic Border Control



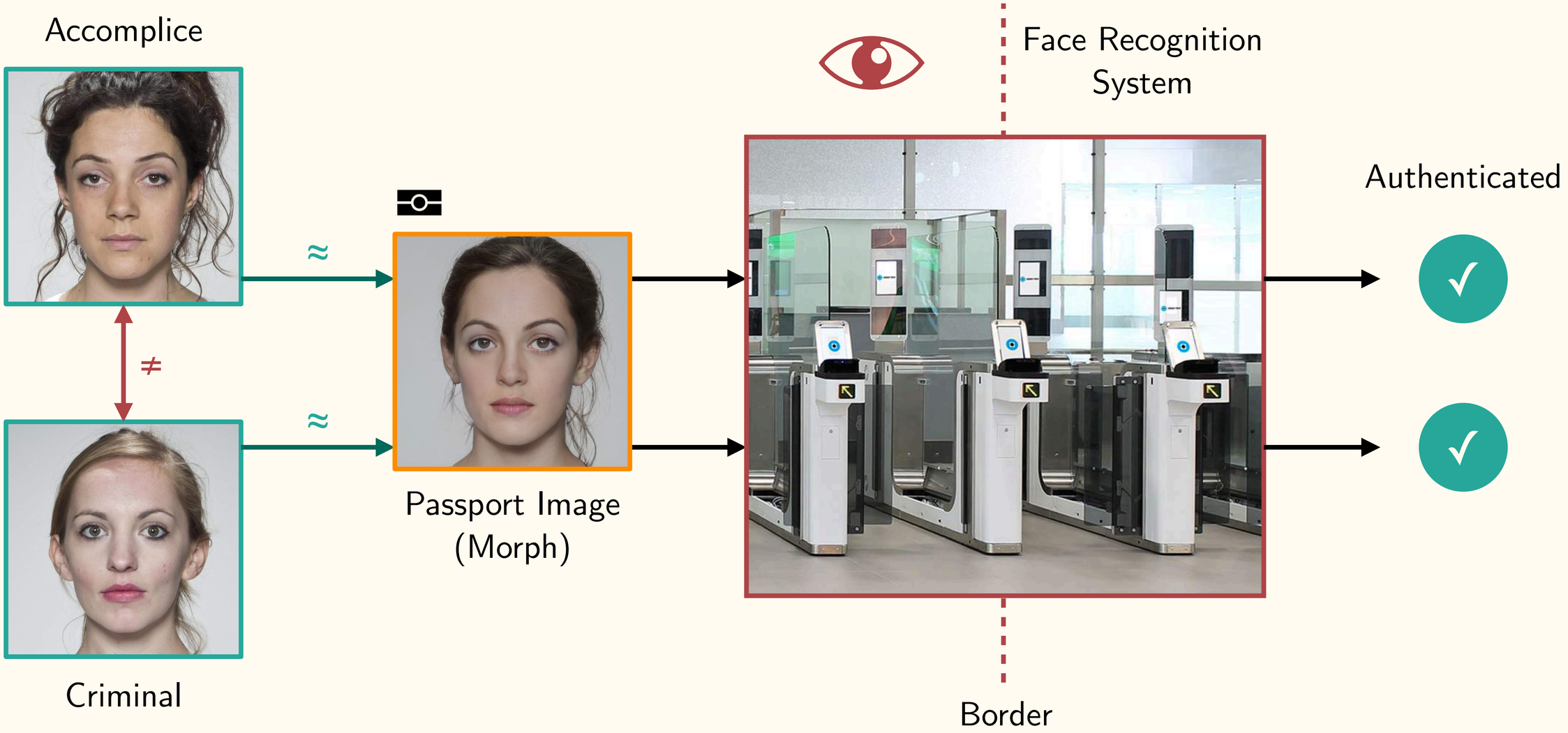
Morphing Attack - Automatic Border Control



Morphing Attack - Automatic Border Control



Morphing Attack - Automatic Border Control



Motivation

Motivation

- Work relating to morphing attacks tends to focus on their detection.

Motivation

- Work relating to morphing attacks tends to focus on their detection.
- Some related issues lack attention:

Motivation

- Work relating to morphing attacks tends to focus on their detection.
- Some related issues lack attention:
  - Little reported on whether latest SOTA face recognition systems remain vulnerable to typical morphing attacks.

Motivation

- Work relating to morphing attacks tends to focus on their detection.
- Some related issues lack attention:
  - Little reported on whether latest SOTA face recognition systems remain vulnerable to typical morphing attacks.
  - Very few public datasets of ‘morphed’ images:

Motivation

- Work relating to morphing attacks tends to focus on their detection.
- Some related issues lack attention:
  - Little reported on whether latest SOTA face recognition systems remain vulnerable to typical morphing attacks.
  - Very few public datasets of 'morphed' images:
    - Advanced Multimedia Security Lab's (AMSL) Face Morph Image dataset

This study

This study

We attempt to fill these gaps by:

This study

We attempt to fill these gaps by:

- Providing new dataset with four different types of morphing attacks generated with original face images from three public face datasets.

This study

We attempt to fill these gaps by:

- Providing new dataset with four different types of morphing attacks generated with original face images from three public face datasets.
- Conducting extensive experiments to assess the vulnerability of SOTA face recognition systems.

This study

We attempt to fill these gaps by:

- Providing new dataset with four different types of morphing attacks generated with original face images from three public face datasets.
- Conducting extensive experiments to assess the vulnerability of SOTA face recognition systems.

Milestone in our understanding of where the current SOTA morph generation algorithms and FR systems are at.

Morph Generation - Datasets

Morph Generation - Datasets

- FERET & FRGCv2

Morph Generation - Datasets

- FERET & FRGCv2
  - Large number of images of different identities

Morph Generation - Datasets

- FERET & FRGCv2
  - Large number of images of different identities
- FRLL

Morph Generation - Datasets

- FERET & FRGCv2
  - Large number of images of different identities
- FRLL
  - Ideal for creating morphing attacks

Morph Generation - Datasets

- FERET & FRGCv2
  - Large number of images of different identities
- FRLL
  - Ideal for creating morphing attacks
  - Close-up frontal face images

Morph Generation - Datasets

- FERET & FRGCv2
  - Large number of images of different identities
- FRLL
  - Ideal for creating morphing attacks
  - Close-up frontal face images
  - 1350 × 1350 resolution

Morph Generation - Datasets

- FERET & FRGCv2
  - Large number of images of different identities
- FRLL
  - Ideal for creating morphing attacks
  - Close-up frontal face images
  - 1350 × 1350 resolution
  - Uniform illumination

Morph Generation - Datasets

- FERET & FRGCv2
  - Large number of images of different identities
- FRLL
  - Ideal for creating morphing attacks
  - Close-up frontal face images
  - 1350 × 1350 resolution
  - Uniform illumination
  - Large varieties in ethnicity, pose, and expression

Morph Generation - Datasets

- FERET & FRGCv2
  - Large number of images of different identities
- FRLL
  - Ideal for creating morphing attacks
  - Close-up frontal face images
  - 1350 × 1350 resolution
  - Uniform illumination
  - Large varieties in ethnicity, pose, and expression
  - Pre-annotated with 189-landmarks

Morph Generation - Tools

Morph Generation - Tools

- Landmark-based morphs:
  - OpenCV
  - FaceMorpher
  - WebMorph
  - Combined Morphs

Morph Generation - Tools

- Landmark-based morphs:
  - OpenCV
  - FaceMorpher
  - WebMorph
  - Combined Morphs
- Generative Adversarial Networks-based morphs:
  - StyleGAN 2

Morph Generation - Landmarks

Identity A

Identity B

Extract Landmarks

Triangulate

Warp

Alpha Blend

Morphed Image

Morph Generation - Landmarks

Identity A

Identity B

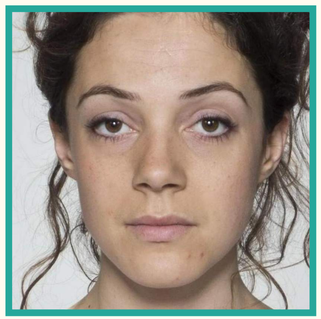
Extract Landmarks

Triangulate

Warp

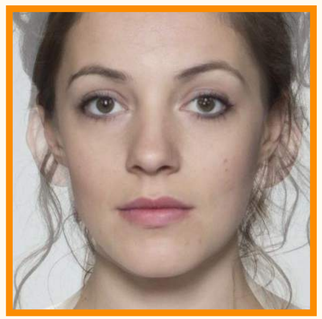
Alpha Blend

Morphed Image



Identity A

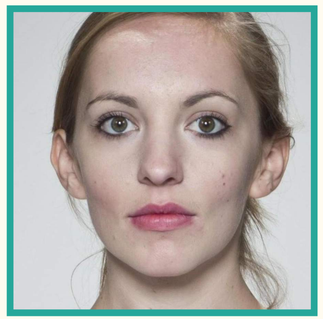
OpenCV



FaceMorpher

WebMorph

Combined Morphs



Identity B

Morph Generation - Landmarks

Identity A

Identity B

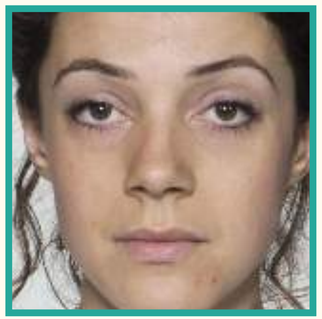
Extract Landmarks

Triangulate

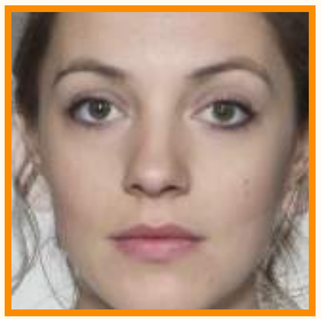
Warp

Alpha Blend

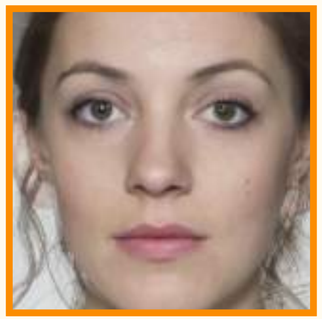
Morphed Image



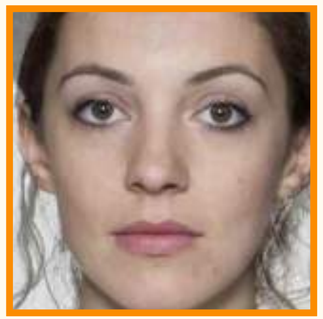
Identity A



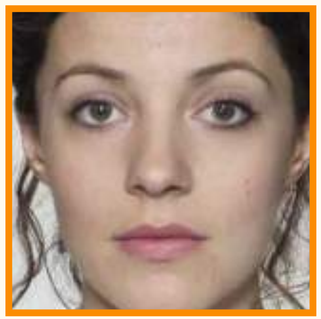
OpenCV



FaceMorpher



WebMorph



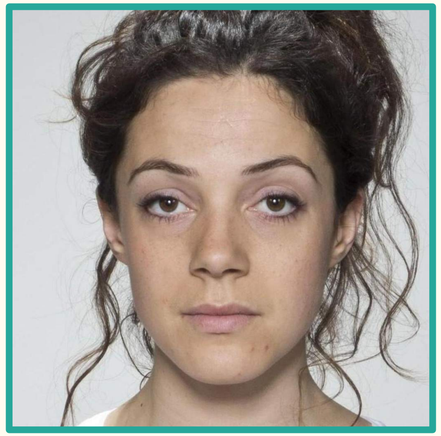
Combined Morphs

Identity B

Morph Generation - StyleGAN 2

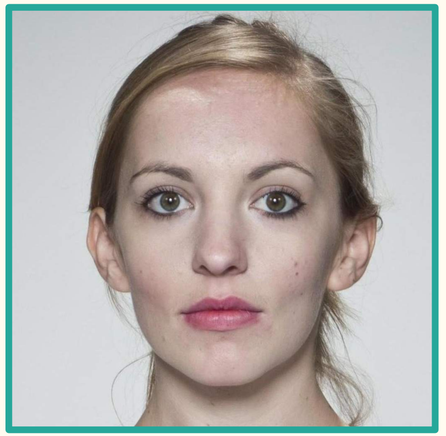
Morph Generation - StyleGAN 2

Identity A



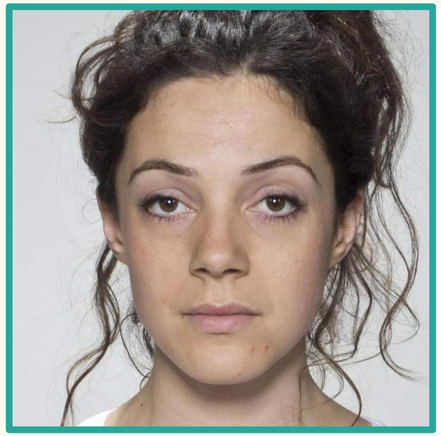
1. Crop source images to FFHQ alignment

Identity B

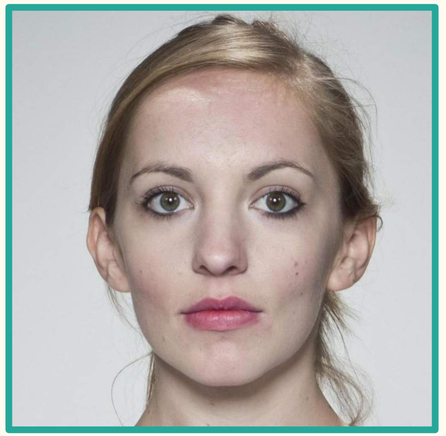


Morph Generation - StyleGAN 2

Identity A



Identity B



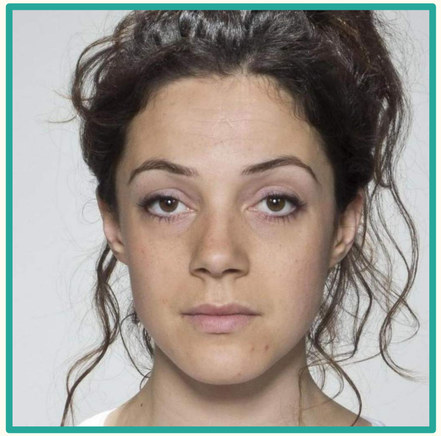
1. Crop source images to FFHQ alignment
2. Project images to StyleGAN's W latent space

Projection A

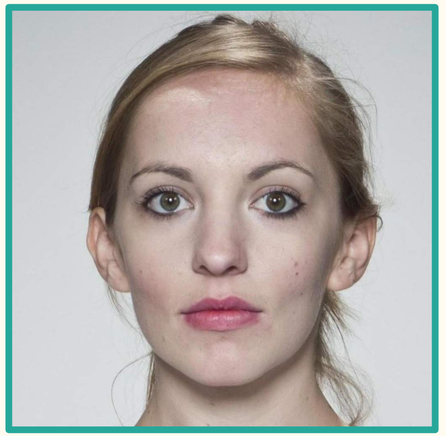
Projection B

Morph Generation - StyleGAN 2

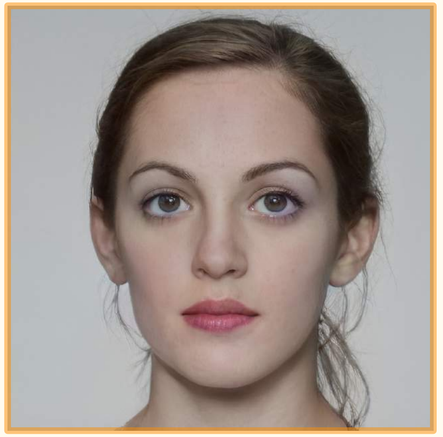
Identity A



Identity B



1. Crop source images to FFHQ alignment
2. Project images to StyleGAN's W latent space
3. Linearly interpolate latent vectors

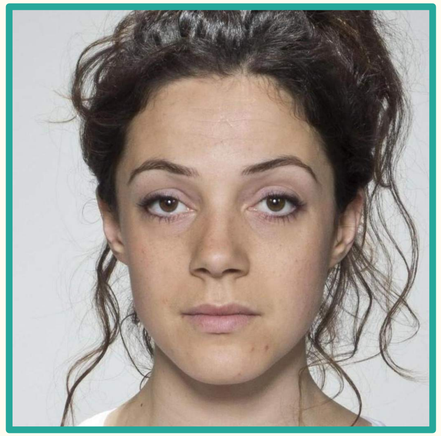


Projection A

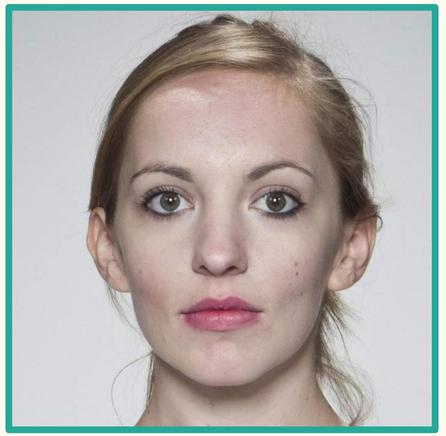
Projection B

Morph Generation - StyleGAN 2

Identity A

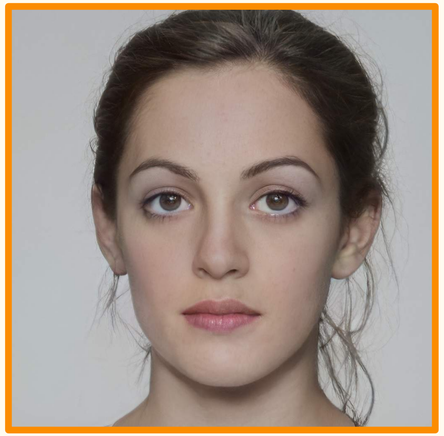


Identity B

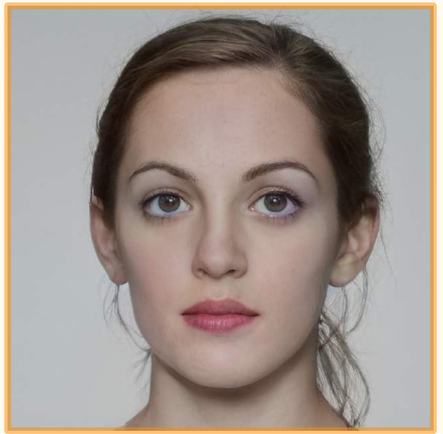


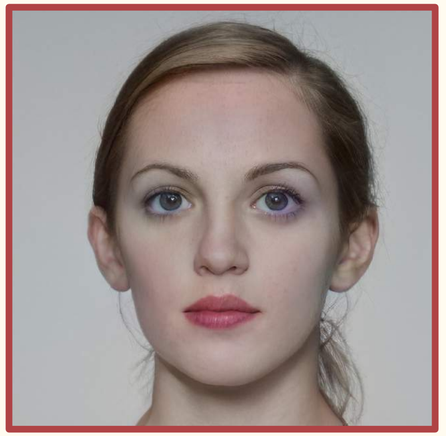
1. Crop source images to FFHQ alignment
2. Project images to StyleGAN's W latent space
3. Linearly interpolate latent vectors
4. Feed interpolated vector back to generator

Projection A



Morph





Projection B

Morph Generation - StyleGAN 2

Identity A

Identity B

1. Crop source images to FFHQ alignment
2. Project images to StyleGAN's W space
3. Linearly interpolate latent vectors
4. Feed interpolated vector back to generator

Projection A

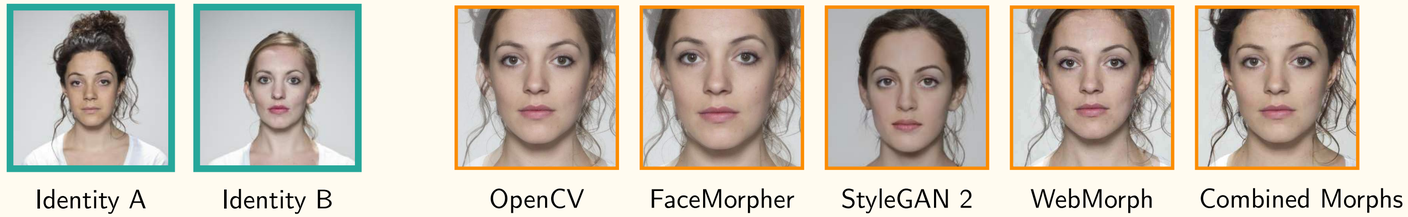
Morph

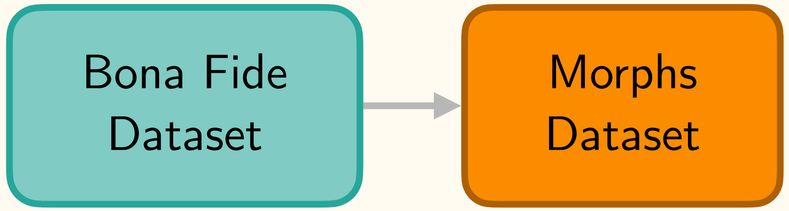
Projection B

- Realistic looking morphs without visual artefacts

Pipeline Summary

Pipeline Summary

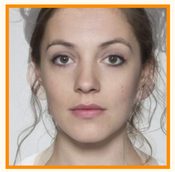


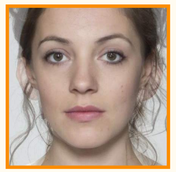


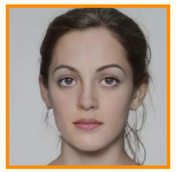
Pipeline Summary

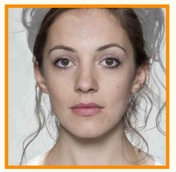

Identity A


Identity B


OpenCV


FaceMorpher


StyleGAN 2


WebMorph


Combined Morphs

Bona Fide Dataset

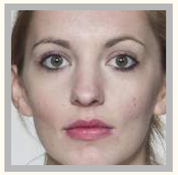
Morphs Dataset

Preprocessing

Embedding Extraction

Cosine Distance

Scores


Reference


Probe

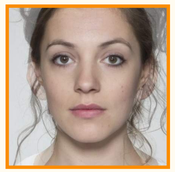
Last FC layers of FRS network when input with the reference or probe image

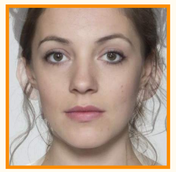
between the 2 embeddings

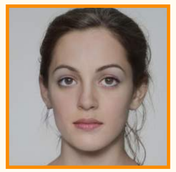
Pipeline Summary

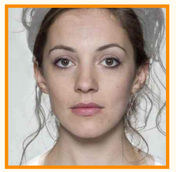

Identity A


Identity B


OpenCV


FaceMorpher


StyleGAN 2


WebMorph


Combined Morphs

Bona Fide Dataset

Morphs Dataset

Preprocessing

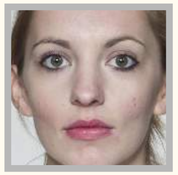
Embedding Extraction

Cosine Distance

Scores

Tune decision threshold

Evaluate Vulnerability of FRS


Reference


Probe

Last FC layers of FRS network when input with the reference or probe image

between the 2 embeddings

Face Recognition Systems (FRS)

Face Recognition Systems (FRS)

- Pre-trained Deep Neural Networks:
  - FaceNet - 99.6%
  - ArcFace - 99.5%
  - VGG-Face - 98.5%

  } Accuracy on LFW dataset

Face Recognition Systems (FRS)

- Pre-trained Deep Neural Networks:
  - FaceNet - 99.6%
  - ArcFace - 99.5%
  - VGG-Face - 98.5%

  } Accuracy on LFW dataset

- Classical Baseline Models:
  - Gabor Jet
  - Inter-Session Variability (ISV) - trained on MOBIO dataset

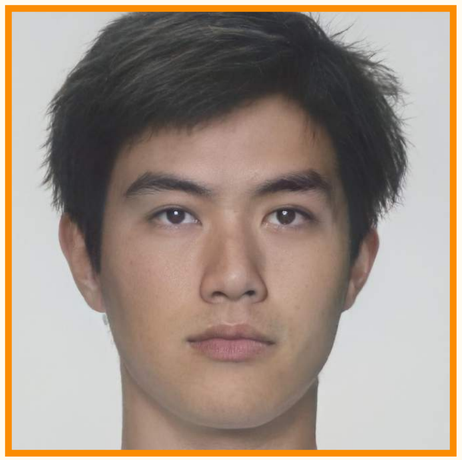
Evaluation Scenarios - Morphing Attack

Evaluation Scenarios - Morphing Attack

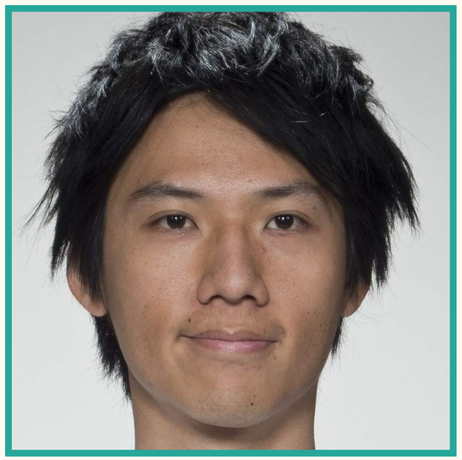
Morphs as **references**:

Evaluation Scenarios - Morphing Attack

Morphs as **references**:



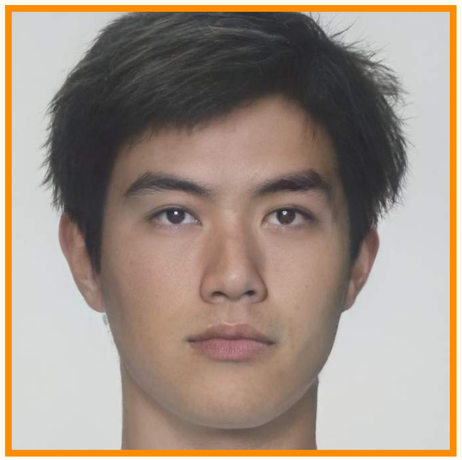
Reference: Neutral MA



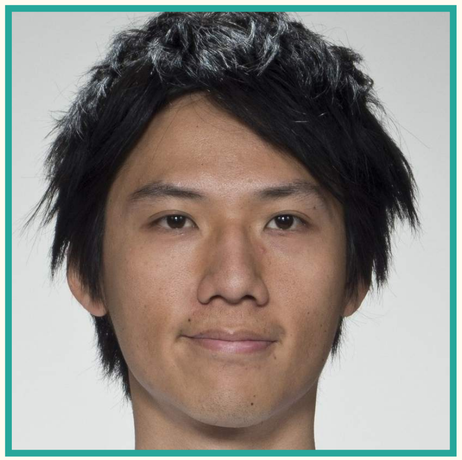
Probe: Smiling BF

Evaluation Scenarios - Morphing Attack

Morphs as **references**:



Reference: Neutral MA

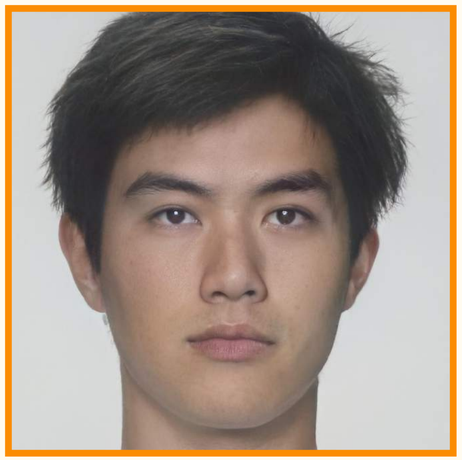


Probe: Smiling BF

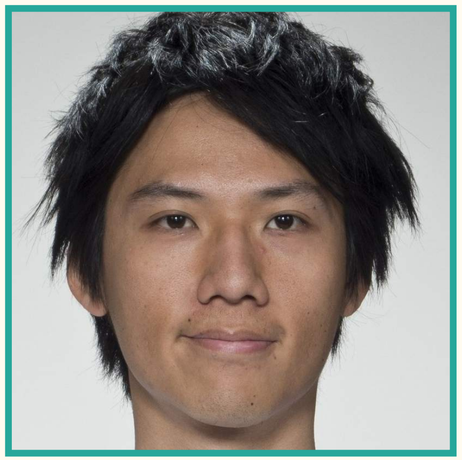
FR system hijacked during enrollment process

Evaluation Scenarios - Morphing Attack

Morphs as **references**:



Reference: Neutral MA



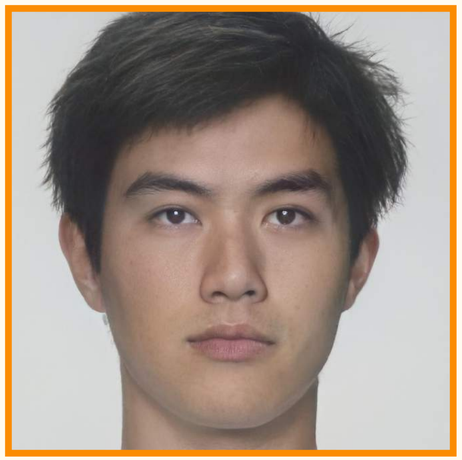
Probe: Smiling BF

FR system hijacked during enrollment process

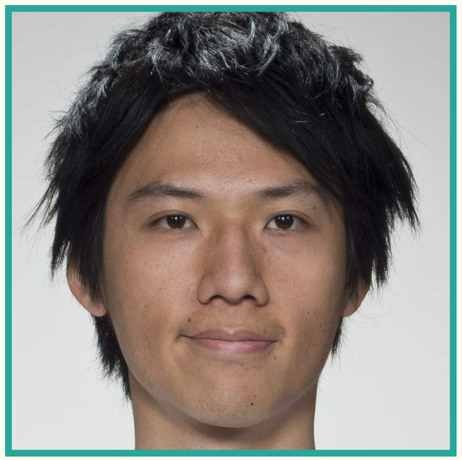
Morphs as **probes**:

Evaluation Scenarios - Morphing Attack

Morphs as **references**:

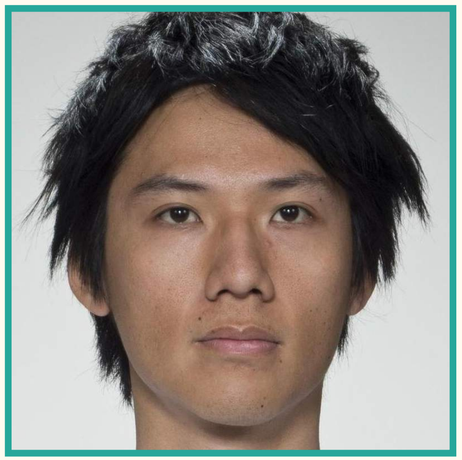


Reference: Neutral MA

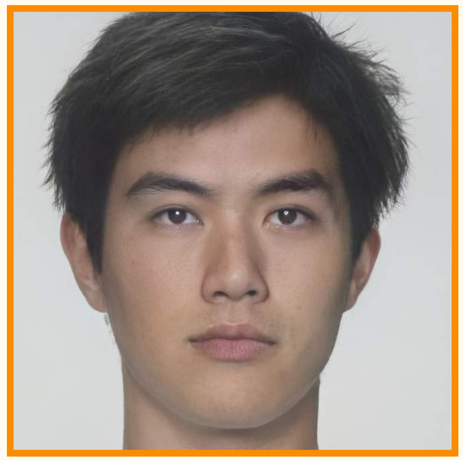


Probe: Smiling BF

Morphs as **probes**:



Reference: Neutral BF

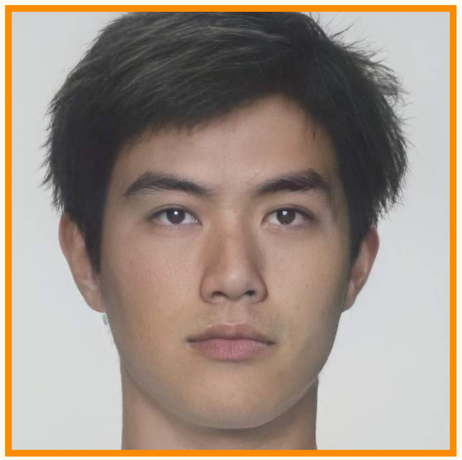


Probe: Neutral MA

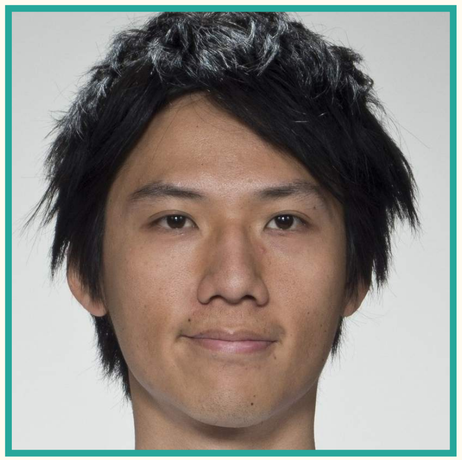
FR system hijacked during enrollment process

Evaluation Scenarios - Morphing Attack

Morphs as **references**:



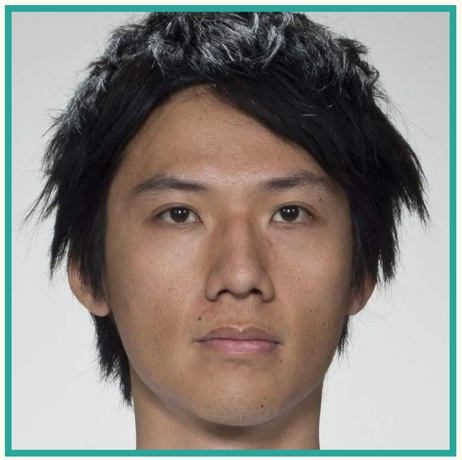
Reference: Neutral MA



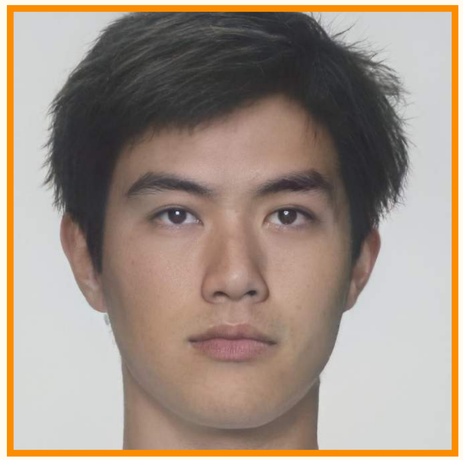
Probe: Smiling BF

FR system hijacked during enrollment process

Morphs as **probes**:



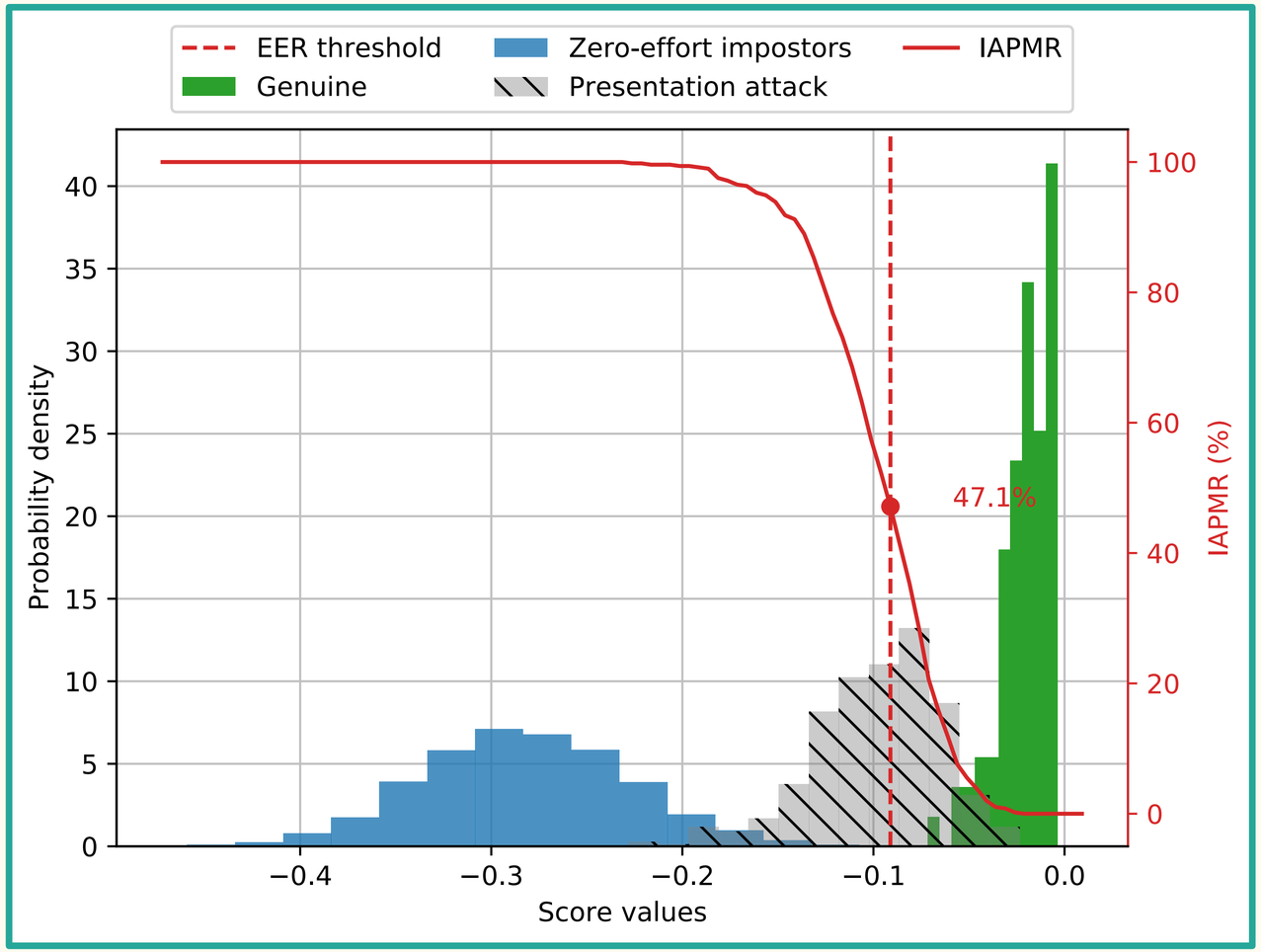
Reference: Neutral BF



Probe: Neutral MA

Similar to presentation attack scenario

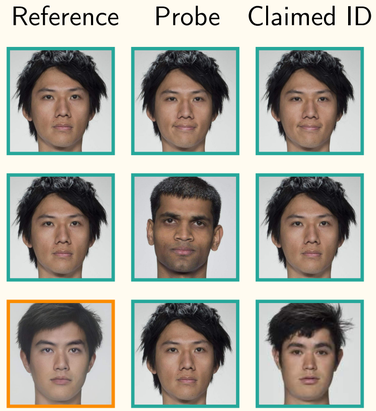
Evaluation Metrics



FRS: VGG, Morphing Tool: OpenCV

Verification Process:

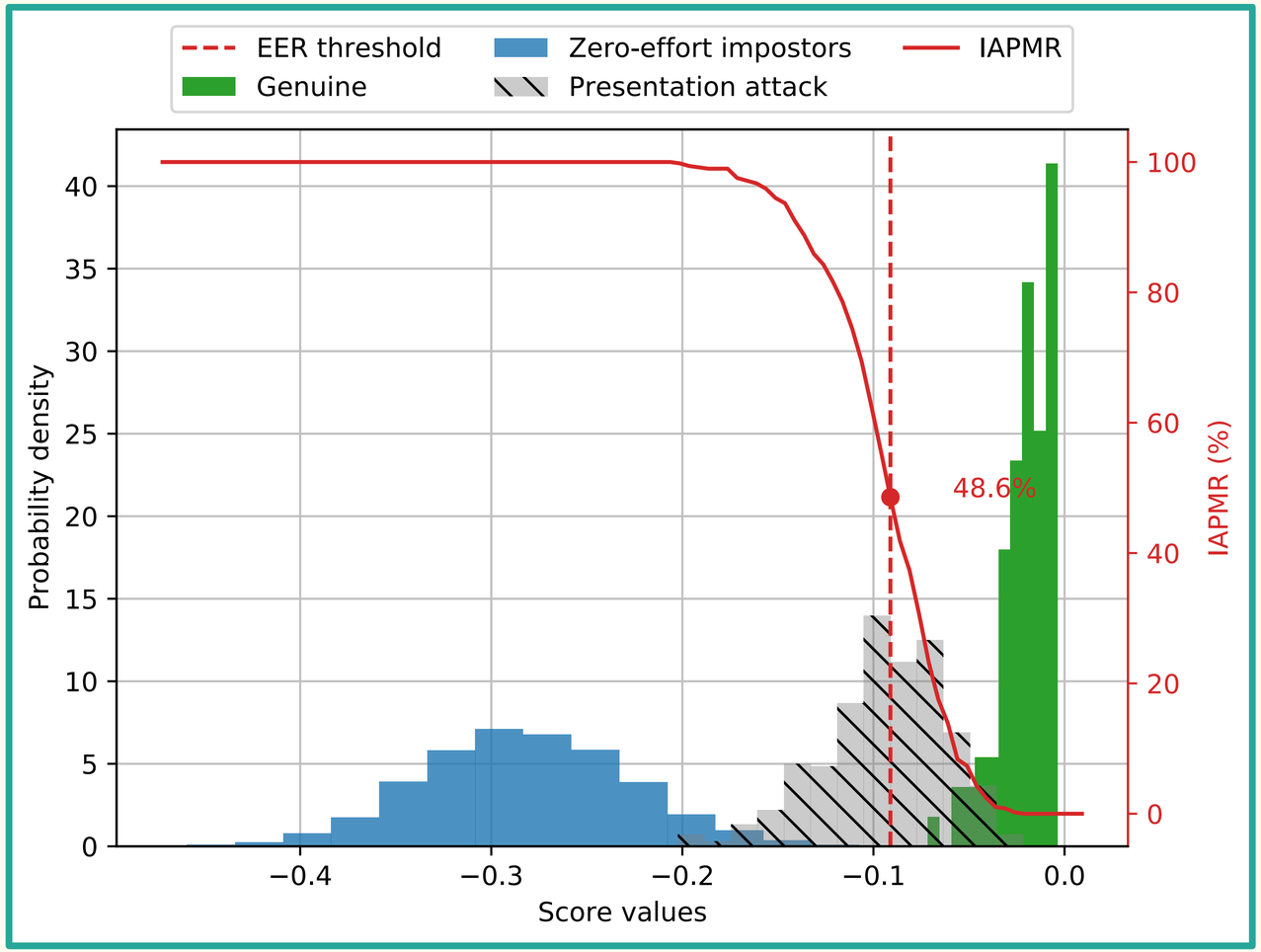
- Genuine User
- Zero-Effort Imposter
- Morph Attack Imposter



Verification Performance:

- False Match Rate (FMR)
- False Non-Match Rate (FNMR)
- Mated-Morph Presentation Match Rate (MMPMR)

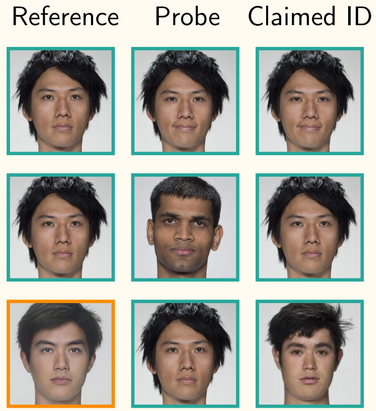
Evaluation Metrics



FRS: VGG, Morphing Tool: FaceMorpher

Verification Process:

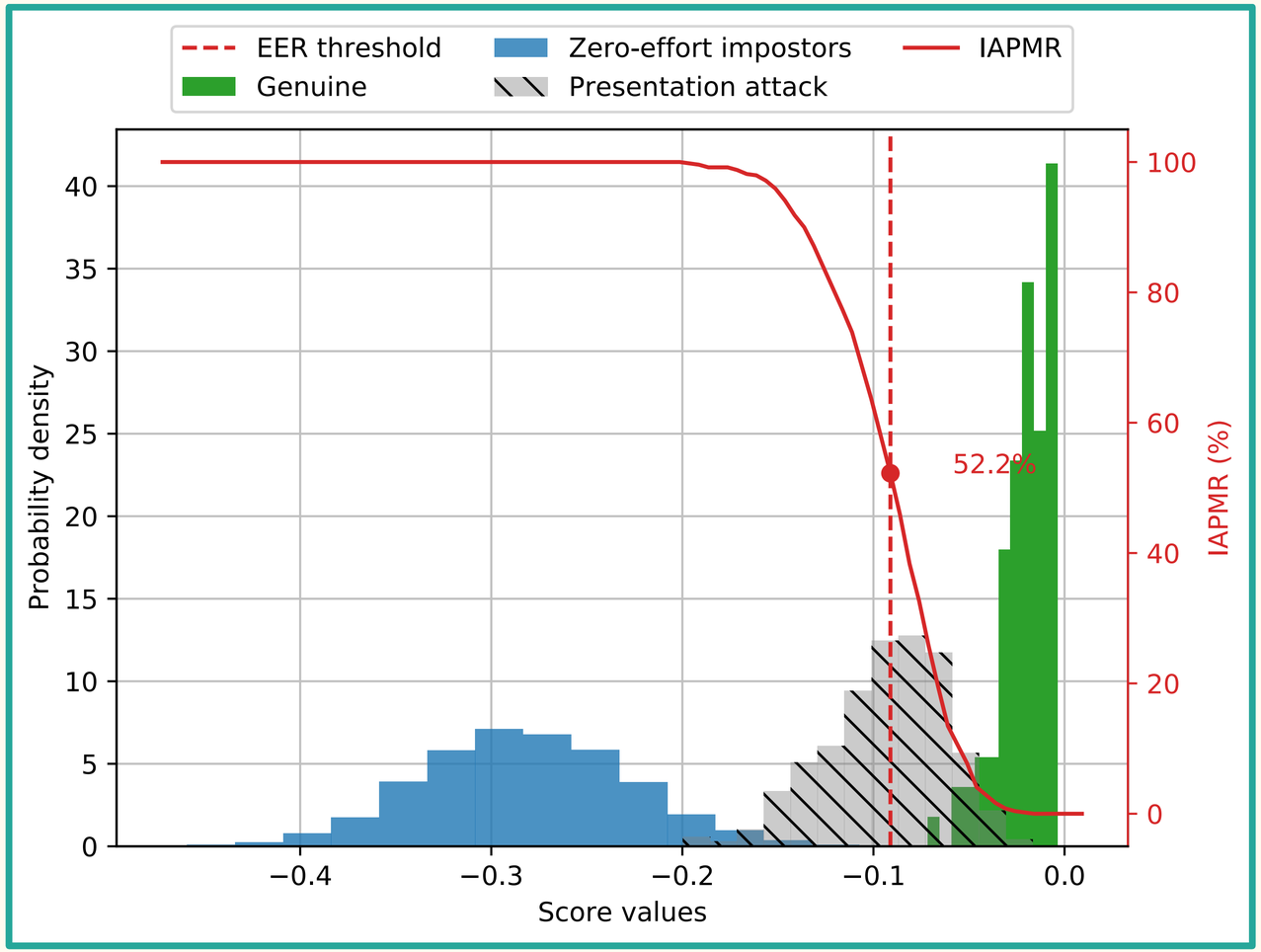
- Genuine User
- Zero-Effort Imposter
- Morph Attack Imposter



Verification Performance:

- False Match Rate (FMR)
- False Non-Match Rate (FNMR)
- Mated-Morph Presentation Match Rate (MMPMR)

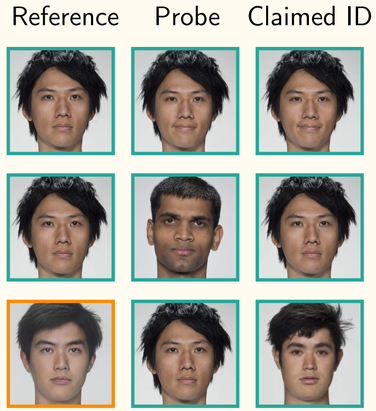
Evaluation Metrics



FRS: VGG, Morphing Tool: WebMorph

Verification Process:

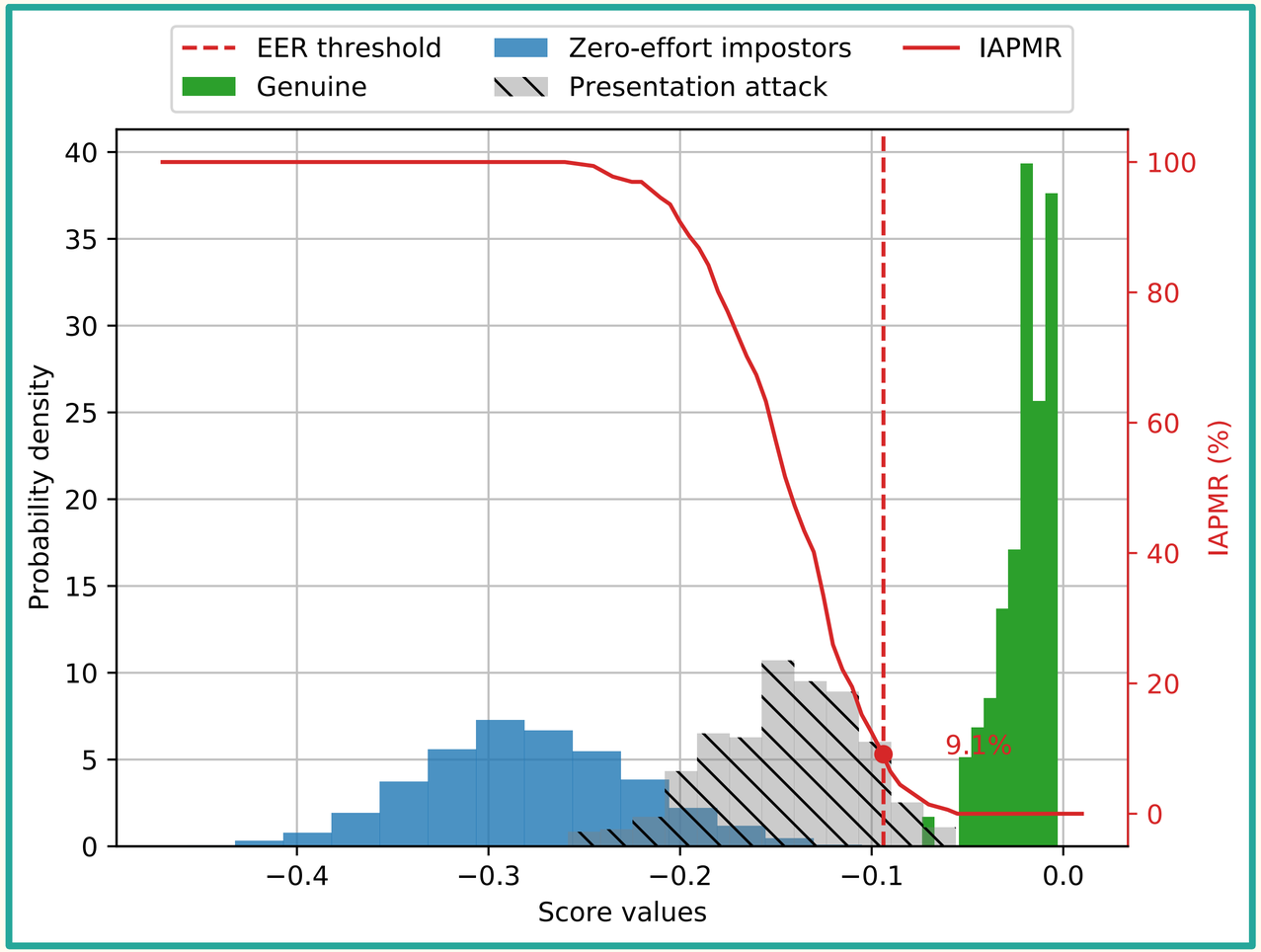
- Genuine User
- Zero-Effort Imposter
- Morph Attack Imposter



Verification Performance:

- False Match Rate (FMR)
- False Non-Match Rate (FNMR)
- Mated-Morph Presentation Match Rate (MMPMR)

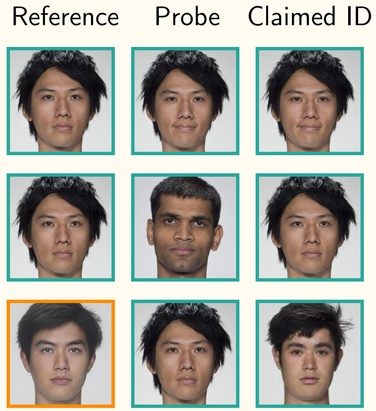
Evaluation Metrics



FRS: VGG, Morphing Tool: StyleGAN 2

Verification Process:

- Genuine User
- Zero-Effort Imposter
- Morph Attack Imposter



Verification Performance:

- False Match Rate (FMR)
- False Non-Match Rate (FNMR)
- Mated-Morph Presentation Match Rate (MMPMR)

Experimental Results

Dataset
FRLL
FERET
FRGC

Experimental Results

Dataset	FRS
FRLL	FaceNet
	ArcFace
	VGG
	Gabor
	ISV
FERET	FaceNet
	ArcFace
	VGG
	Gabor
	ISV
FRGC	FaceNet
	ArcFace
	VGG
	Gabor
	ISV

Experimental Results

Dataset	FRS	OpenCV	FaceMorpher	StyleGAN2	WebMorph	AMSL
FRLL	FaceNet					
	ArcFace					
	VGG					
	Gabor					
	ISV					
FERET	FaceNet					
	ArcFace					
	VGG					
	Gabor					
	ISV					
FRGC	FaceNet					
	ArcFace					
	VGG					
	Gabor					
	ISV					

Experimental Results

Dataset	FRS	OpenCV	FaceMorpher	StyleGAN2	WebMorph	AMSL
FRLL	FaceNet					
	ArcFace					
	VGG					
	Gabor					
	ISV					
FERET	FaceNet				N/A	N/A
	ArcFace				N/A	N/A
	VGG				N/A	N/A
	Gabor				N/A	N/A
	ISV				N/A	N/A
FRGC	FaceNet				N/A	N/A
	ArcFace				N/A	N/A
	VGG				N/A	N/A
	Gabor				N/A	N/A
	ISV				N/A	N/A

Experimental Results

MMPMR @ FMR = 0.1%

Dataset	FRS	OpenCV	FaceMorpher	StyleGAN2	WebMorph	AMSL
FRLL	FaceNet					
	ArcFace					
	VGG					
	Gabor					
	ISV					
FERET	FaceNet				N/A	N/A
	ArcFace				N/A	N/A
	VGG				N/A	N/A
	Gabor				N/A	N/A
	ISV				N/A	N/A
FRGC	FaceNet				N/A	N/A
	ArcFace				N/A	N/A
	VGG				N/A	N/A
	Gabor				N/A	N/A
	ISV				N/A	N/A

Experimental Results

MMPMR @ FMR = 0.1% (morphs as references — morphs as probes) [%]

Dataset	FRS	OpenCV	FaceMorpher	StyleGAN2	WebMorph	AMSL
FRLL	FaceNet	83.3 — 72.0	64.5 — 68.2	5.9 — 11.0	82.7 — 70.8	89.2 — 92.5
	ArcFace	59.8 — 73.8	57.6 — 75.3	9.8 — 18.3	60.9 — 73.8	58.0 — 79.4
	VGG	39.7 — 48.6	23.4 — 47.1	3.0 — 9.1	38.2 — 52.2	65.7 — 89.8
	Gabor	87.2 — 100.0	83.9 — 99.4	11.8 — 37.9	85.4 — 100.0	86.3 — 99.9
	ISV	59.8 — 97.8	56.1 — 96.1	9.2 — 43.6	59.5 — 97.4	55.3 — 99.9
FERET	FaceNet	41.1 — 40.6	39.9 — 40.3	1.6 — 1.3	N/A	N/A
	ArcFace	34.6 — 35.2	34.1 — 34.8	2.4 — 2.5	N/A	N/A
	VGG	22.0 — 21.0	20.5 — 18.3	2.0 — 1.5	N/A	N/A
	Gabor	66.6 — 90.9	63.7 — 88.5	1.3 — 40.8	N/A	N/A
	ISV	44.8 — 58.4	42.6 — 56.5	2.7 — 3.4	N/A	N/A
FRGC	FaceNet	6.9 — 5.9	7.0 — 5.7	1.0 — 0.7	N/A	N/A
	ArcFace	11.9 — 10.8	12.1 — 11.2	0.5 — 0.4	N/A	N/A
	VGG	5.5 — 4.5	5.1 — 4.8	0.7 — 0.4	N/A	N/A
	Gabor	7.1 — 80.8	6.7 — 81.0	0.6 — 75.8	N/A	N/A
	ISV	4.2 — 6.5	3.5 — 6.2	0.6 — 0.6	N/A	N/A

Higher score indicates higher vulnerability

Experimental Results

MMPMR @ FMR = 0.1% (morphs as references — morphs as probes) [%]

Dataset	FRS	OpenCV	FaceMorpher	StyleGAN2	WebMorph	AMSL
FRLL	**FaceNet**	**83.3 — 72.0**	**64.5 — 68.2**	**5.9 — 11.0**	**82.7 — 70.8**	**89.2 — 92.5**
	ArcFace	59.8 — 73.8	57.6 — 75.3	9.8 — 18.3	60.9 — 73.8	58.0 — 79.4
	VGG	**39.7 — 48.6**	**23.4 — 47.1**	**3.0 — 9.1**	**38.2 — 52.2**	**65.7 — 89.8**
	Gabor	87.2 — 100.0	83.9 — 99.4	11.8 — 37.9	85.4 — 100.0	86.3 — 99.9
	ISV	59.8 — 97.8	56.1 — 96.1	9.2 — 43.6	59.5 — 97.4	55.3 — 99.9
FERET	**FaceNet**	**41.1 — 40.6**	**39.9 — 40.3**	**1.6 — 1.3**	N/A	N/A
	ArcFace	34.6 — 35.2	34.1 — 34.8	2.4 — 2.5	N/A	N/A
	VGG	**22.0 — 21.0**	**20.5 — 18.3**	**2.0 — 1.5**	N/A	N/A
	Gabor	66.6 — 90.9	63.7 — 88.5	1.3 — 40.8	N/A	N/A
	ISV	44.8 — 58.4	42.6 — 56.5	2.7 — 3.4	N/A	N/A
FRGC	**FaceNet**	**6.9 — 5.9**	**7.0 — 5.7**	**1.0 — 0.7**	N/A	N/A
	ArcFace	11.9 — 10.8	12.1 — 11.2	0.5 — 0.4	N/A	N/A
	VGG	**5.5 — 4.5**	**5.1 — 4.8**	**0.7 — 0.4**	N/A	N/A
	Gabor	7.1 — 80.8	6.7 — 81.0	0.6 — 75.8	N/A	N/A
	ISV	4.2 — 6.5	3.5 — 6.2	0.6 — 0.6	N/A	N/A

Higher score indicates higher vulnerability

Experimental Results

MMPMR @ FMR = 0.1% (morphs as references — morphs as probes) [%]

Dataset	FRS	OpenCV	FaceMorpher	StyleGAN2	WebMorph	AMSL
FRLL	FaceNet	83.3 — 72.0	64.5 — 68.2	**5.9 — 11.0**	82.7 — 70.8	89.2 — 92.5
	ArcFace	59.8 — 73.8	57.6 — 75.3	**9.8 — 18.3**	60.9 — 73.8	58.0 — 79.4
	VGG	39.7 — 48.6	23.4 — 47.1	**3.0 — 9.1**	38.2 — 52.2	65.7 — 89.8
	Gabor	87.2 — 100.0	83.9 — 99.4	**11.8 — 37.9**	85.4 — 100.0	86.3 — 99.9
	ISV	59.8 — 97.8	56.1 — 96.1	**9.2 — 43.6**	59.5 — 97.4	55.3 — 99.9
FERET	FaceNet	41.1 — 40.6	39.9 — 40.3	**1.6 — 1.3**	N/A	N/A
	ArcFace	34.6 — 35.2	34.1 — 34.8	**2.4 — 2.5**	N/A	N/A
	VGG	22.0 — 21.0	20.5 — 18.3	**2.0 — 1.5**	N/A	N/A
	Gabor	66.6 — 90.9	63.7 — 88.5	**1.3 — 40.8**	N/A	N/A
	ISV	44.8 — 58.4	42.6 — 56.5	**2.7 — 3.4**	N/A	N/A
FRGC	FaceNet	6.9 — 5.9	7.0 — 5.7	**1.0 — 0.7**	N/A	N/A
	ArcFace	11.9 — 10.8	12.1 — 11.2	**0.5 — 0.4**	N/A	N/A
	VGG	5.5 — 4.5	5.1 — 4.8	**0.7 — 0.4**	N/A	N/A
	Gabor	7.1 — 80.8	6.7 — 81.0	**0.6 — 75.8**	N/A	N/A
	ISV	4.2 — 6.5	3.5 — 6.2	**0.6 — 0.6**	N/A	N/A

Higher score indicates higher vulnerability

Observations

Observations

- More accurate face recognition system → more vulnerable to morphing attacks
  - In line with observations made for presentation attacks.
  - Especially evident when comparing FaceNet with VGG-Face.
  - Regardless of whether used as references or probes.

Observations

- More accurate face recognition system → more vulnerable to morphing attacks
  - In line with observations made for presentation attacks.
  - Especially evident when comparing FaceNet with VGG-Face.
  - Regardless of whether used as references or probes.
- StyleGAN 2 morphs do **not** pose significant threats to SOTA recognition systems.
  - Slight increase in MMPMR for FRLL morphs → high quality original images may lead to slightly more accurate morphs.

Conclusion

- Conducted extensive vulnerability assessments (5 recognition systems on 3 image databases with 5 different morphing attacks in 2 different scenarios).
- An accurate face recognition system FaceNet is more vulnerable to the morphing attacks than others.
- GAN-based morphs do not yet pose a significant threat to modern recognition systems.
  - Additional identity loss to ensure both source bona fide identities are *preserved* in their projections → the final morphs may become more threatening.

Thank you !

Room 207-2, Idiap Research Institute

www.idiap.ch/~esarkar/

+41 78 82 50 754

eklavya.sarkar@idiap.ch