Are GAN-based Morphs threatening Face Recognition?

Eklavya SARKAR
Research Assistant, Idiap Research Institute and EPFL
Content

- Problem
- Data Generation
- Experiments and Evaluation Protocols
- Results
- Summary
Problem
Problem

Morphing Attack: When two individuals’ face images is combined into a single ‘morphed’ image using a morphing algorithm.
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.
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.
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
Morphing Attack - Automatic Border Control
Morphing Attack - Automatic Border Control

Accomplice

Criminal
Morphing Attack - Automatic Border Control

Accomplice ≠ Criminal
Morphing Attack - Automatic Border Control

Accomplice ≠ ≈ Passport Image (Morph)

Criminal
Morphing Attack - Automatic Border Control

Accomplice ≠ ≈ ≈ Passport Image (Morph)

Criminal
Morphing Attack - Automatic Border Control

Accomplice

≠

≈

Criminal

Passport Image (Morph)

Face Recognition System

Border
Morphing Attack - Automatic Border Control

Accomplice

Criminal

Passport Image (Morph)

Face Recognition System

Authenticated

Border
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:
  ▪ No clear understanding on whether the latest FR systems are vulnerable to both ‘classical’ and latest GAN-based morphing attacks.
Motivation

- Work relating to morphing attacks tends to focus on their detection.
- Some related issues lack attention:
  - No clear understanding on whether the latest FR systems are vulnerable to both ‘classical’ and latest GAN-based 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:
  ▪ No clear understanding on whether the latest FR systems are vulnerable to both ‘classical’ and latest GAN-based morphing attacks.
  ▪ Very few public datasets of morphed images.
  ▪ Modern morphing techniques rarely publicly released.
Motivation

- Work relating to morphing attacks tends to focus on their detection.
- Some related issues lack attention:
  - No clear understanding on whether the latest FR systems are vulnerable to both ‘classical’ and latest GAN-based morphing attacks.
  - Very few public datasets of morphed images.
  - Modern morphing techniques rarely publicly released.
  - Lack of evaluation protocols.
Contributions
Contributions

This paper provides the following three contributions:
Contributions

This paper provides the following three contributions:

- Provide an open source morphing tool\(^1\) for generating morphing attacks.

\(^1\)https://gitlab.idiap.ch/bob/bob.paper.icassp2022_morph_generate
Contributions

This paper provides the following three contributions:

- Provide an open source morphing tool\(^1\) for generating morphing attacks.
- Providing new datasets with morphed images generated using different algorithms on two public face datasets.

\(^1\)https://gitlab.idiap.ch/bob/bob.paper.icassp2022_morph_generate
Contributions

This paper provides the following three contributions:

● Provide an open source morphing tool\(^1\) for generating morphing attacks.

● Providing new datasets with morphed images generated using different algorithms on two public face datasets.

● Conducting extensive experiments to assess the vulnerability of SOTA face recognition systems.

\(^1\)https://gitlab.idiap.ch/bob/bob.paper.icassp2022_morph_generate
Morph Generation - Tools
Morph Generation - Tools

Traditional: **Landmark based morphs**

- OpenCV
- FaceMorpher
Morph Generation - Tools

Traditional: Landmark based morphs

- OpenCV
- FaceMorpher

Modern: GAN based morphs

- StyleGAN 2
- MIPGAN-II
Morph Generation - Landmarks

Identity A → Extract Landmarks → Triangulate → Warp → Alpha Blend → Morphed Image

Identity B
Morph Generation - Landmarks

Identity A
Extract Landmarks
Triangulate
Warp
Alpha Blend
Morphed Image

Identity B

Identity A
OpenCV
FaceMorpher
Identity B
Morph Generation - Landmarks

Identity A
Extract Landmarks
Identity B
Triangulate
Warp
Alpha Blend
Morphed Image

Identity A
OpenCV
FaceMorpher
Identity B
Morph Generation - StyleGAN 2
Morph Generation - StyleGAN 2

1. Crop source images to FFHQ alignment
Morph Generation - StyleGAN 2

1. Crop source images to FFHQ alignment
2. Project images to StyleGAN’s W latent space
Morph Generation - StyleGAN 2

1. Crop source images to FFHQ alignment
2. Project images to StyleGAN’s $W$ latent space
3. Linearly interpolate latent vectors
Morph Generation - StyleGAN 2

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
Morph Generation - MIPGAN II

Zhang, H., Venkatesh, S., Ramachandra, R., Raja, K., Damer, N. and Busch, C., 2021. Mipgan—generating strong and high quality morphing attacks using identity prior driven GAN. IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(3), pp.365-383.
Morph Generation - MIPGAN II

- *Optimises the latent vector* of the StyleGAN morph
  - To improve the perceptual fidelity, quality, identity factor of the StyleGAN morph.

Zhang, H., Venkatesh, S., Ramachandra, R., Raja, K., Damer, N. and Busch, C., 2021. Mipgan—generating strong and high quality morphing attacks using identity prior driven GAN. IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(3), pp.365-383.
Morph Generation - MIPGAN II

- Optimises the latent vector of the StyleGAN morph
  - To improve the perceptual fidelity, quality, identity factor of the StyleGAN morph.
- The weighted sum of 3 additional losses are used:

Zhang, H., Venkatesh, S., Ramachandra, R., Raja, K., Damer, N. and Busch, C., 2021. Mipgan—generating strong and high quality morphing attacks using identity prior driven GAN. IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(3), pp.365-383.
Morph Generation - MIPGAN II

- **Optimises the latent vector** of the StyleGAN morph
  - To improve the perceptual fidelity, quality, identity factor of the StyleGAN morph.

- The weighted sum of 3 additional losses are used:
  - $L_1$ Perceptual loss: maintains visual fidelity.
Morph Generation - MIPGAN II

- **Optimises the latent vector** of the StyleGAN morph
  - To improve the perceptual fidelity, quality, identity factor of the StyleGAN morph.

- The weighted sum of **3 additional losses** are used:
  - $L_1$ Perceptual loss: maintains visual fidelity.
  - $L_2$ Identity loss: conserves identity of input images.

---

Zhang, H., Venkatesh, S., Ramachandra, R., Raja, K., Damer, N. and Busch, C., 2021. Mipgan—generating strong and high quality morphing attacks using identity prior driven GAN. IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(3), pp.365-383.
Morph Generation - MIPGAN II

- **Optimises** the *latent vector* of the StyleGAN morph
  - To improve the perceptual fidelity, quality, identity factor of the StyleGAN morph.
- The weighted sum of **3 additional losses** are used:
  - $L_1$ Perceptual loss: maintains visual fidelity.
  - $L_2$ Identity loss: conserves identity of input images.
  - $L_3$ ID-Difference: equally balances between the input images.

Zhang, H., Venkatesh, S., Ramachandra, R., Raja, K., Damer, N. and Busch, C., 2021. Mipgan—generating strong and high quality morphing attacks using identity prior driven GAN. IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(3), pp.365-383.
Morph Generation - MIPGAN II

- *Optimises* the *latent vector* of the StyleGAN morph
  - To improve the perceptual fidelity, quality, identity factor of the StyleGAN morph.

- The weighted sum of *3 additional losses* are used:
  - $\mathcal{L}_1$ Perceptual loss: maintains visual fidelity.
  - $\mathcal{L}_2$ Identity loss: conserves identity of input images.
  - $\mathcal{L}_3$ ID-Difference: equally balances between the input images.
  - $\mathcal{L}_4$ MS-SSIM: improves structural visibility.

---

Zhang, H., Venkatesh, S., Ramachandra, R., Raja, K., Damer, N. and Busch, C., 2021. Mipgan—generating strong and high quality morphing attacks using identity prior driven GAN. IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(3), pp.365-383.
Morph Generation - MIPGAN II

- **Optimises** the *latent vector* of the StyleGAN morph
  - To improve the perceptual fidelity, quality, identity factor of the StyleGAN morph.

- The weighted sum of 3 additional losses are used:
  - $\mathcal{L}_1$ **Perceptual loss**: maintains visual fidelity.
  - $\mathcal{L}_2$ **Identity loss**: conserves identity of input images.
  - $\mathcal{L}_3$ **ID-Difference**: equally balances between the input images.
  - $\mathcal{L}_4$ **MS-SSIM**: improves structural visibility.

\[
\mathcal{L} = \lambda_1 \mathcal{L}_1 + \lambda_2 \mathcal{L}_2 + \lambda_3 \mathcal{L}_3 + \lambda_4 \mathcal{L}_4
\]
Morph Generation - MIPGAN II

Step 0

StyleGAN2
Morph

Optimization through $\mathcal{L}$

Step 150

MIPGAN-II
Morph
Experiments
Pipeline Summary
Pipeline Summary

Identity A

Identity B

Bona Fide Dataset
Pipeline Summary

Bona Fide Dataset → Morphs Dataset

Identity A → Identity B

OpenCV → FaceMorpher → StyleGAN 2 → MIPGAN-II
Pipeline Summary

Bona Fide Dataset

Morphs Dataset

Preprocessing

Reference

Probe
Pipeline Summary

Bona Fide Dataset

OpenCV

FaceMorpher

StyleGAN 2

MIPGAN-II

Morphs Dataset

Preprocessing

Embedding Extraction

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

Reference

Probe
Pipeline Summary

Bona Fide Dataset

Morphs Dataset

Preprocessing

Embedding Extraction

Cosine Distance

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

between the 2 embeddings
Pipeline Summary

Bona Fide Dataset

Morphs Dataset

Preprocessing

Embedding Extraction

Cosine Distance

Scores

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

between the 2 embeddings

Identity A
(100,400),(175,600)

Identity B
(200,400),(275,600)

OpenCV

FaceMorpher

StyleGAN 2

MIPGAN-II

Reference

Probe
Pipeline Summary

Bona Fide Dataset → Morphs Dataset

Preprocessing → Embedding Extraction → Cosine Distance → Scores

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

between the 2 embeddings

Reference → Probe

Tune decision threshold

Evaluate Vulnerability of FRS

Scores
Evaluation and Metrics

FRS: VGG, Morphing Tool: OpenCV
Evaluation and Metrics

Verification Process:

FRS: VGG, Morphing Tool: OpenCV
Evaluation and Metrics

Verification Process:
- Genuine User

FRS: VGG, Morphing Tool: OpenCV
Evaluation and Metrics

Verification Process:
- Genuine User
- Zero-Effort Imposter

FRS: VGG, Morphing Tool: OpenCV
Evaluation and Metrics

Verification Process:
- Genuine User
- Zero-Effort Imposter
- Morph Attack Imposter

FRS: VGG, Morphing Tool: OpenCV
Evaluation and Metrics

Verification Process:
- Genuine User
- Zero-Effort Imposter
- Morph Attack Imposter

Verification Performance:

FRS: VGG, Morphing Tool: OpenCV
Evaluation and Metrics

Verification Process:
- Genuine User
- Zero-Effort Imposter
- Morph Attack Imposter

Verification Performance:
- Mated-Morph Presentation Match Rate — (MMPMR [%])

FRS: VGG, Morphing Tool: OpenCV
Evaluation and Metrics

Verification Process:
- Genuine User
- Zero-Effort Imposter
- Morph Attack Imposter

Verification Performance:
- Mated-Morph Presentation Match Rate — (MMPMR [%])

FRS: VGG, Morphing Tool: FaceMorpher
Evaluation and Metrics

Verification Process:
- Genuine User
- Zero-Effort Imposter
- Morph Attack Imposter

Verification Performance:
- Mated-Morph Presentation Match Rate — (MMPMR [%])

FRS: VGG, Morphing Tool: WebMorph
Evaluation and Metrics

Verification Process:
- Genuine User
- Zero-Effort Imposter
- Morph Attack Imposter

Verification Performance:
- Mated-Morph Presentation Match Rate — (MMPMR [%])

FRS: VGG, Morphing Tool: StyleGAN 2
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%

- Classical Baseline Models:
  - Inter-Session Variability (ISV) - trained on MOBIO dataset
Morph Generation - Datasets
Morph Generation - Datasets

- FERET
- FRLL
Morph Generation - Datasets

- FERET
- FRLL
  - Close-up frontal face images
Morph Generation - Datasets

- FERET
- FRLL

- Close-up frontal face images
- 1350 × 1350 resolution
Morph Generation - Datasets

- FERET
- FRLL
  - Close-up frontal face images
  - 1350 × 1350 resolution
  - Uniform illumination
Morph Generation - Datasets

- FERET
- FRLL
  - Close-up frontal face images
  - 1350 × 1350 resolution
  - Uniform illumination
  - Large varieties in ethnicity, pose, and expression
Evaluation Scenarios - Morphing Attack
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**  
Probes: **Smiling BF**

Morphs as **probes**: Reference: **Neutral BF**  
Probes: **Neutral MA**

FR system hijacked during enrollment process

Similar to presentation attack scenario
## Experimental Results

| Tool         |  |  |  |
|--------------|---|---|---|
| OpenCV       |  |  |  |
| FaceMorpher  |  |  |  |
| StyleGAN2    |  |  |  |
| MIPGAN-II    |  |  |  |
## Experimental Results

| Tool     | FRS          |
|----------|--------------|
| OpenCV   | FaceNet, ArcFace, VGG, ISV |
| FaceMorpher | FaceNet, ArcFace, VGG, ISV |
| StyleGAN2 | FaceNet, ArcFace, VGG, ISV |
| MIPGAN-II | FaceNet, ArcFace, VGG, ISV |
## Experimental Results

MMPMR @ FMR = 0.1%

| Tool         | FRS        |
|--------------|------------|
| OpenCV       | FaceNet    |
|              | ArcFace    |
|              | VGG        |
|              | ISV        |
| FaceMorpher  | FaceNet    |
|              | ArcFace    |
|              | VGG        |
|              | ISV        |
| StyleGAN2    | FaceNet    |
|              | ArcFace    |
|              | VGG        |
|              | ISV        |
| MIPGAN-II    | FaceNet    |
|              | ArcFace    |
|              | VGG        |
|              | ISV        |
## Experimental Results

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

| Tool       | FRS  |
|------------|------|
| OpenCV     | FaceNet, ArcFace, VGG, ISV |
| FaceMorpher| FaceNet, ArcFace, VGG, ISV |
| StyleGAN2  | FaceNet, ArcFace, VGG, ISV |
| MIPGAN-II  | FaceNet, ArcFace, VGG, ISV |
## Experimental Results

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

| Tool       | FRS       |
|------------|-----------|
| **OpenCV** | FaceNet   |
|            | ArcFace   |
|            | VGG       |
|            | ISV       |
| **FaceMorpher** | FaceNet   |
|            | ArcFace   |
|            | VGG       |
|            | ISV       |
| **StyleGAN2** | FaceNet   |
|            | ArcFace   |
|            | VGG       |
|            | ISV       |
| **MIPGAN-II** | FaceNet   |
|            | ArcFace   |
|            | VGG       |
|            | ISV       |

Higher score indicates higher vulnerability
## Experimental Results

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

| Tool        | 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   |             |

Higher score indicates higher vulnerability
## Experimental Results

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

| Tool       | Tool       | FRS   | FRLL   | FERET  |
|------------|------------|-------|--------|--------|
|            | FaceNet    | 83.3  | 72.0   | 41.1   |
|            | ArcFace    | 59.8  | 73.8   | 34.6   |
|            | VGG        | 39.7  | 48.6   | 22.0   |
|            | ISV        | 59.8  | 97.8   | 44.8   |
| OpenCV     |            | 83.3  | 72.0   | 41.1   |
|            | FaceNet    | 64.5  | 68.2   | 39.9   |
|            | ArcFace    | 57.6  | 75.3   | 34.1   |
|            | VGG        | 23.4  | 47.1   | 20.5   |
|            | ISV        | 56.1  | 96.1   | 42.6   |
| FaceMorpher|            | 64.5  | 68.2   | 39.9   |
|            | FaceNet    | 57.6  | 75.3   | 34.1   |
|            | ArcFace    | 23.4  | 47.1   | 20.5   |
|            | ISV        | 56.1  | 96.1   | 42.6   |
|            | FaceNet    | 59.8  | 97.8   | 44.8   |
|            | ArcFace    | 59.8  | 97.8   | 44.8   |
|            | VGG        | 39.7  | 48.6   | 22.0   |
|            | ISV        | 59.8  | 97.8   | 44.8   |
| StyleGAN2  |            | 5.9   | 11.0   | 1.6    |
|            | ArcFace    | 9.8   | 18.3   | 2.4    |
|            | VGG        | 3.0   | 9.1    | 2.0    |
|            | ISV        | 9.2   | 43.6   | 2.7    |
|            | FaceNet    | 47.2  | 62.7   | 32.9   |
|            | ArcFace    | 32.0  | 46.5   | 26.0   |
|            | VGG        | 15.9  | 30.4   | 14.5   |
|            | ISV        | 3.6   | 23.7   | 7.3    |
| MIPGAN-II  |            | 47.2  | 62.7   | 32.9   |
|            | ArcFace    | 32.0  | 46.5   | 26.0   |
|            | VGG        | 15.9  | 30.4   | 14.5   |
|            | ISV        | 3.6   | 23.7   | 7.3    |

Higher score indicates higher vulnerability
Analysis
Analysis

- StyleGAN 2 morphs do not pose a significant threat to SOTA FR recognition systems (compared to landmark-based morphs).
Analysis

- StyleGAN 2 morphs do **not** pose a significant threat to SOTA FR recognition systems (compared to landmark-based morphs).
  - Likely because of the *original pixels* of both images still present in the features after the landmark-based morphs pipeline is applied.
Analysis

- StyleGAN 2 morphs do not pose a significant threat to SOTA FR recognition systems (compared to landmark-based morphs).
  - Likely because of the original pixels of both images still present in the features after the landmark-based morphs pipeline is applied.
  - Conversely, StyleGAN converses no pixel traces or identity features of the original subjects.
Analysis

- StyleGAN 2 morphs do not pose a significant threat to SOTA FR recognition systems (compared to landmark-based morphs).
  - Likely because of the original pixels of both images still present in the features after the landmark-based morphs pipeline is applied.
  - Conversely, StyleGAN converses no pixel traces or identity features of the original subjects.
  - The synthesised morphed image is instead is perceived as a new, different identity altogether.
Analysis

- StyleGAN 2 morphs do **not** pose a significant threat to SOTA FR recognition systems (compared to landmark-based morphs).
  - Likely because of the original pixels of both images still present in the features after the landmark-based morphs pipeline is applied.
  - Conversely, StyleGAN converses no pixel traces or identity features of the original subjects.
  - The synthesised morphed image is instead is perceived as a new, different identity altogether.
- MIPGAN-II morphs which use extra losses to conserve identity are more threatening.
Summary
Summary

- Generated different types of morphs, and conducted extensive face recognition vulnerability assessments.
Summary

- Generated different types of morphs, and conducted extensive face recognition vulnerability assessments.

- Results show that ‘classical’ morphs are still more of a threat than GAN-based ones, despite their higher visual quality.
Summary

- Generated different types of morphs, and conducted extensive face recognition vulnerability assessments.

- Results show that ‘classical’ morphs are still more of a threat than GAN-based ones, despite their higher visual quality.

➡ We publicly release:

- Open-source morphing tool.
- Generated morph datasets.
- Package for running vulnerability experiments.
Thank you!

Idiap Research Institute

www.idiap.ch/~esarkar/

+ 41 27 72 06 322

eklavya.sarkar@idiap.ch