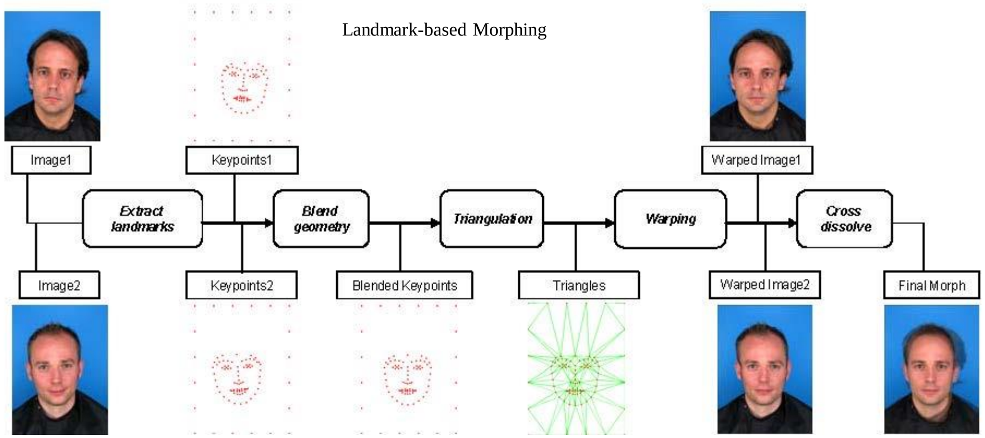
Face Morph Generation and Attack Detection

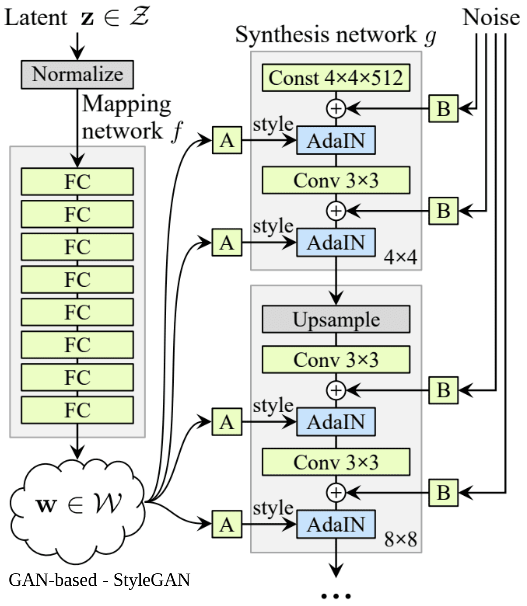
Center for Identification Technology Research (CITeR)

Nasser Nasrabadi (WVU), Chen Liu (CU), Zander Blasingame (CU) and David Doermann (UB)

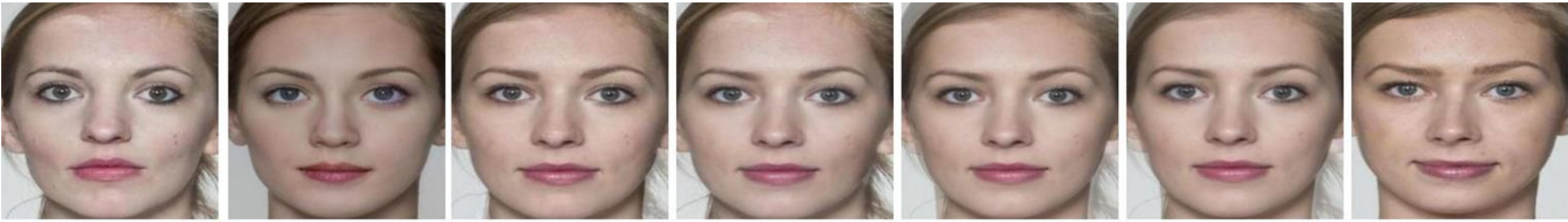
Morph Generation

- Two major methods of morph generation in the literature:
  1. Landmark-based morphing (e.g., OpenCV, FaceMorpher, Combined, and Splicing)
  2. GAN-based (deep learning) morphing (e.g., MorGAN, StyleGAN, and MIPGAN)





GAN-based - StyleGAN


(a) Bonafide #1 (b) StyleGAN2 (c) WebMorph (d) Combined Morph (e) OpenCV (f) Facemorpher (g) Bonafide #2

WVU Wavelet-based landmark morph generation

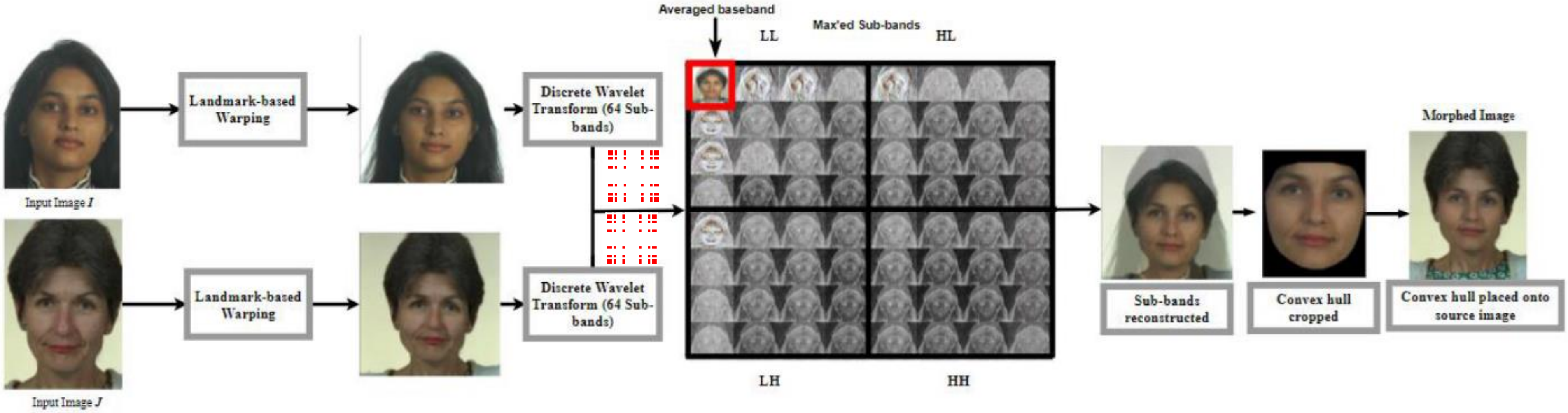
- Replaces alpha-blending stage
- 3-level uniform wavelet decomposition

Alpha Blending

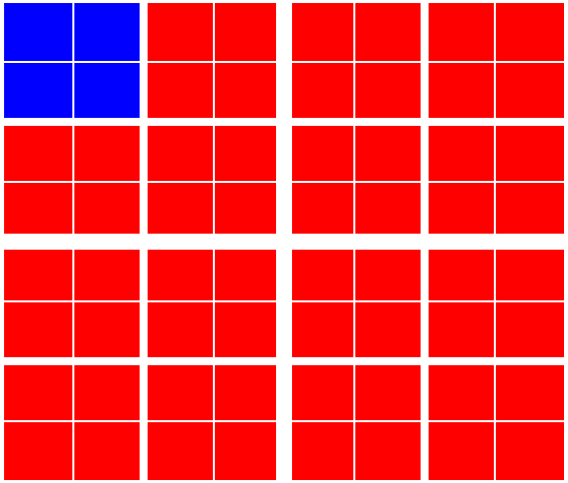
$$I_\alpha(\mathbf{p}) = (1-\alpha)\cdot I_0\left(w_{P_\alpha \to P_0}(\mathbf{p})\right) + \alpha \cdot I_1\left(w_{P_\alpha \to P_1}(\mathbf{p})\right)$$

where:

- $\mathbf{p}$ is a generic pixel position;
- α is the frame weight factor;

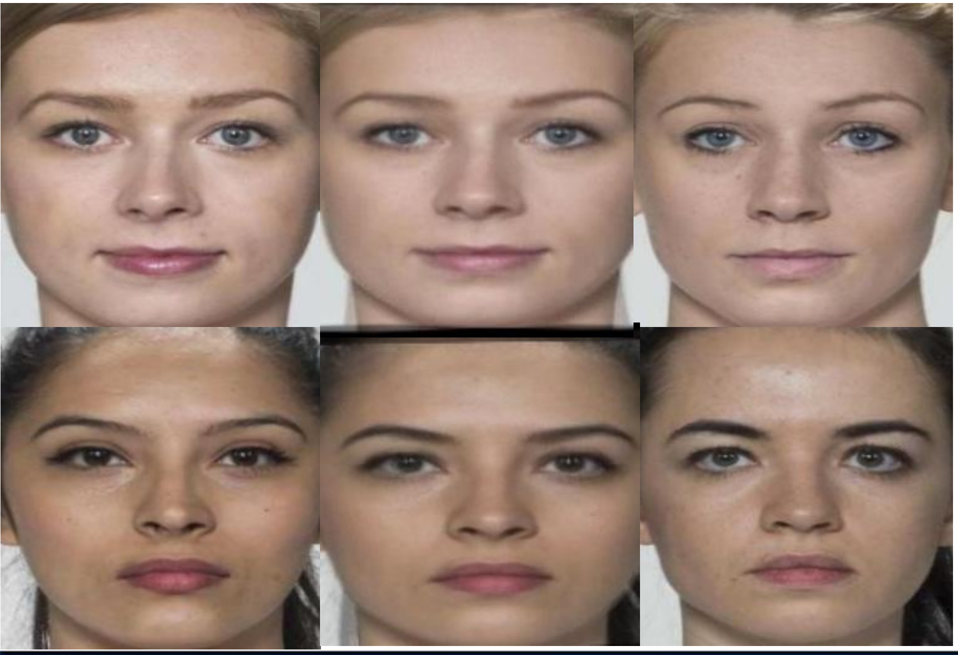


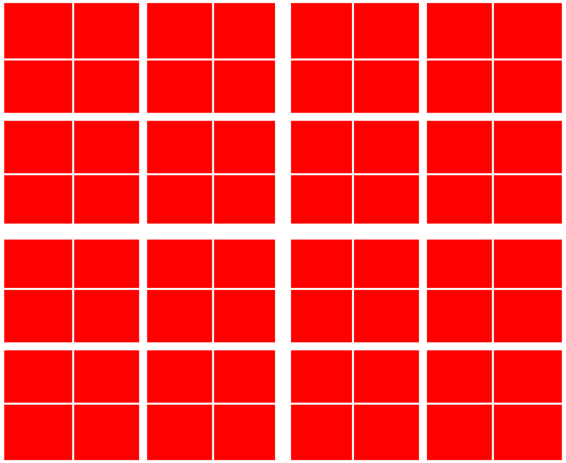
Wavelet-based Fusion



Colors:
Mean
Max

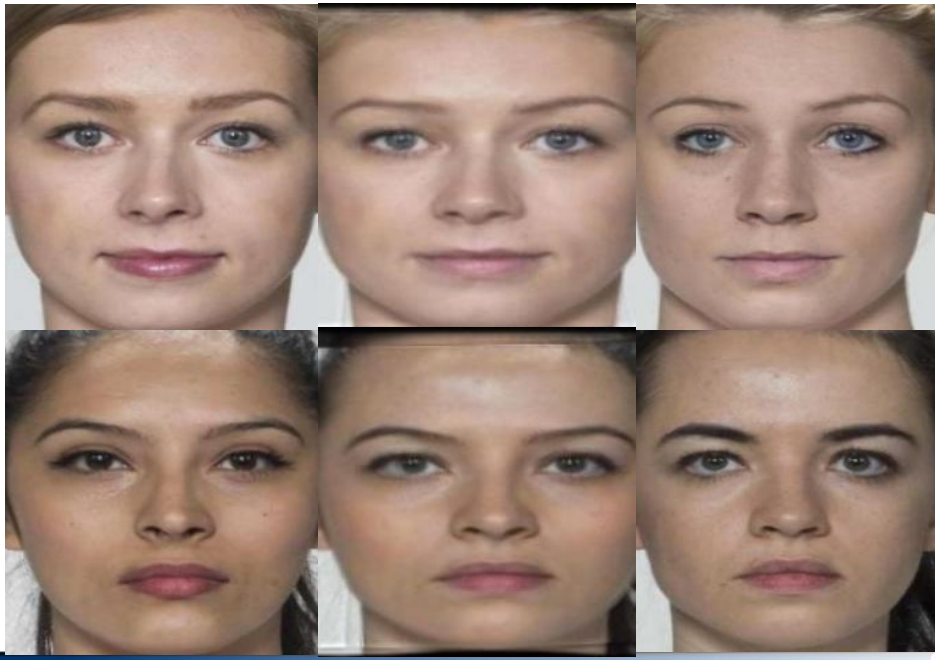
Subject A | Wavelet-based Morphed Image | Subject B





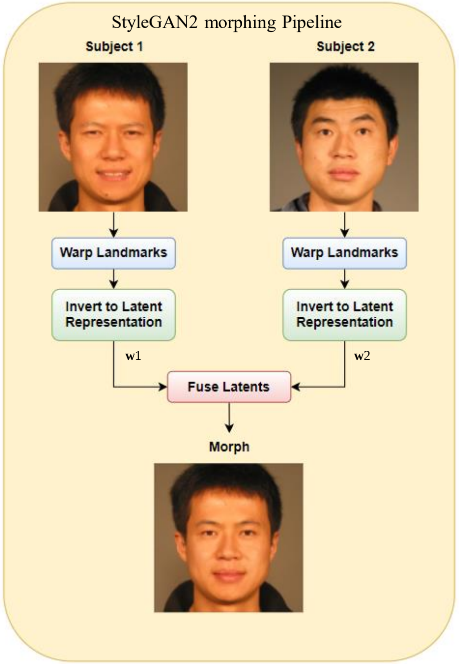
Colors:
Max

Subject A | Wavelet-based Morphed Image | Subject B

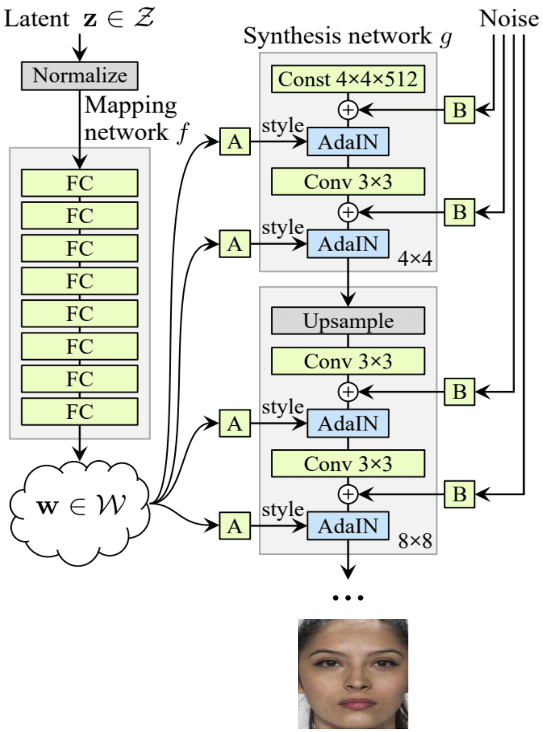


StyleGAN-Based Morph Image Generation with Warp Landmarks

- Utilizes a pre-trained StyleGAN2 deep network.
- Subject 1 and 2 faces are warped to a common landmarks.
- Images are inverted into latent space.
- The latent codes of two subjects (**w**1 , **w**2) are combined (**w**1+**w**2) and then decoded into a morph image.
- Issues retaining identity of bona fide subjects.

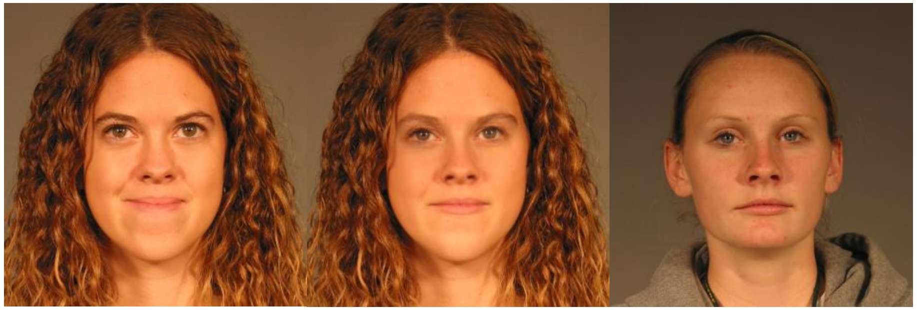

StyleGAN2 morphing Pipeline
Subject 1
Subject 2
Warp Landmarks
Warp Landmarks
Invert to Latent Representation
Invert to Latent Representation
w1
w2
Fuse Latents
Morph

StyleGAN



Examples of WVU FRGC OpenCV, FaceMorpher, StyleGAN2 and Wavelet-based Morphed Faces

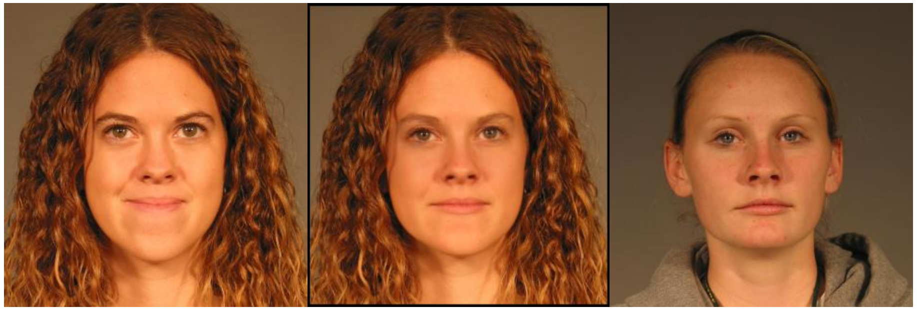
- Morphed face region (convex region) is cropped and blended onto subject A background.



Subject A

Morphed image using OpenCV

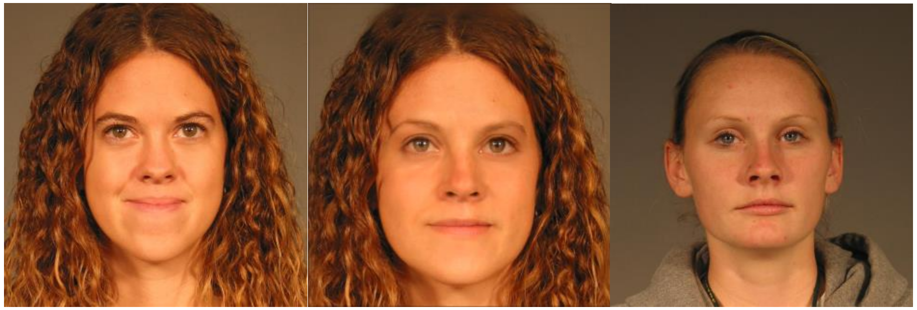
Subject B



Subject A

Morphed image using FaceMorpher

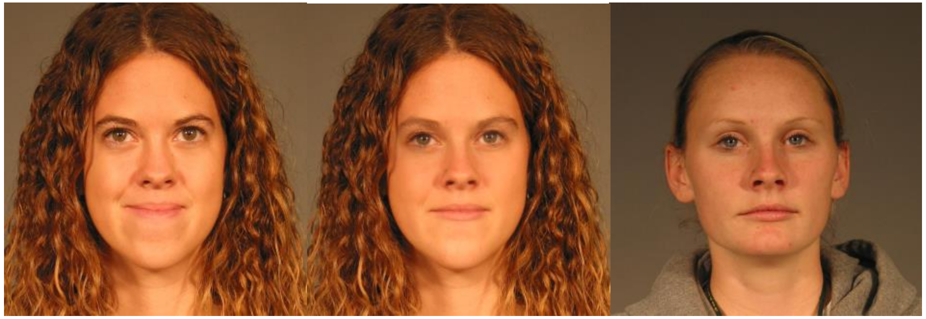
Subject B



Subject A

Morphed image using SyleGAN2

Subject B



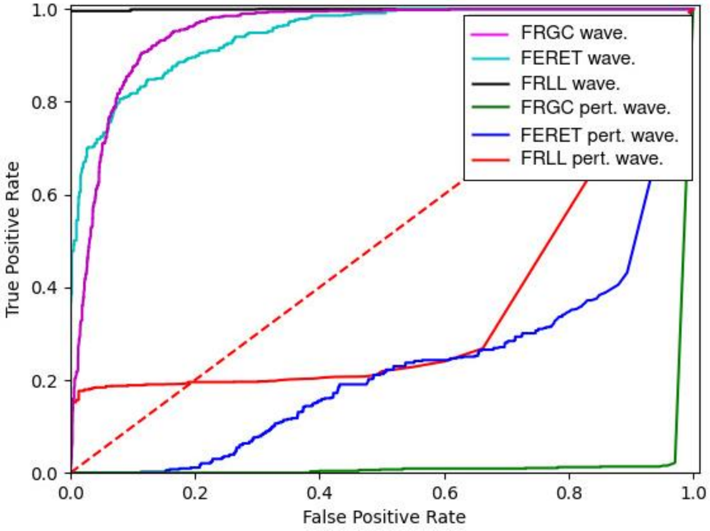
Subject A

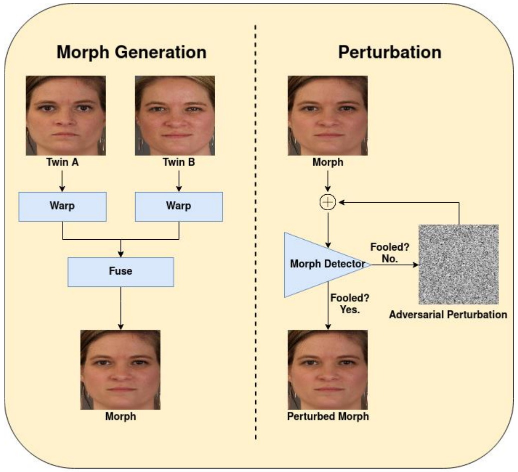
Morphed image using wavelet-based

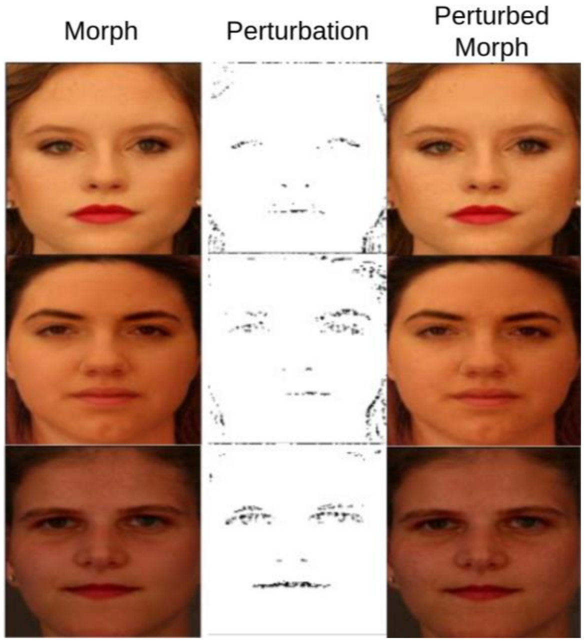
Subject B

Adversarial Perturbation

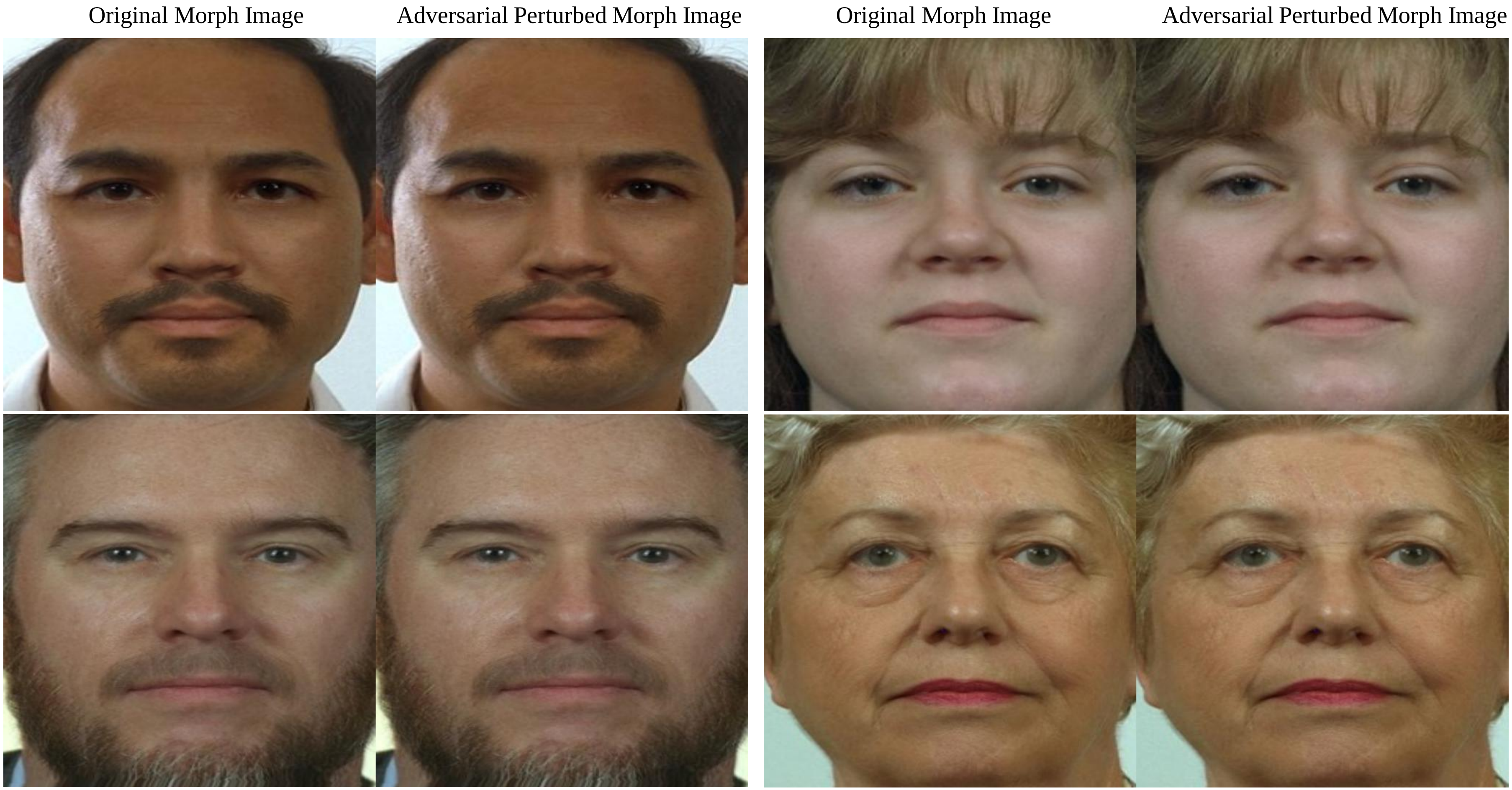
- Structurally significant noise is added to morphed image pixels to erroneously classify a morph as a genuine.
- Uses a Binary Iterative Method (BIM) to perturb images while clipping pixel values at a maximum distance to make noise visually imperceivable.
- Imperceptible adversarial perturbations were added to the wavelet-based morphed images with the intention of fooling our deep morph detector.
- ROC curves show the performance of WVU detector on the original and adversarially perturbed morphed faces for the WVU wavelet-based FRGC, FERET, and FRLL datasets.







Adversarially Perturbed Morphs



Original Morph Image

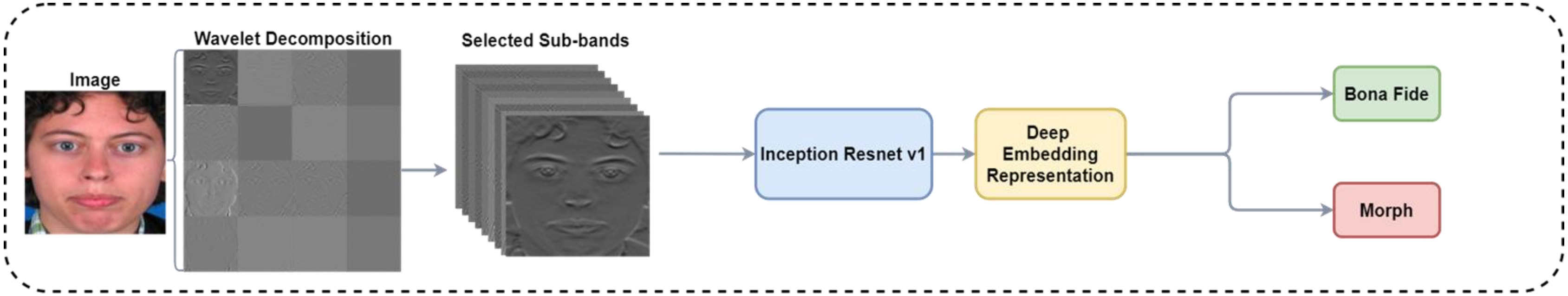
Adversarial Perturbed Morph Image

Original Morph Image

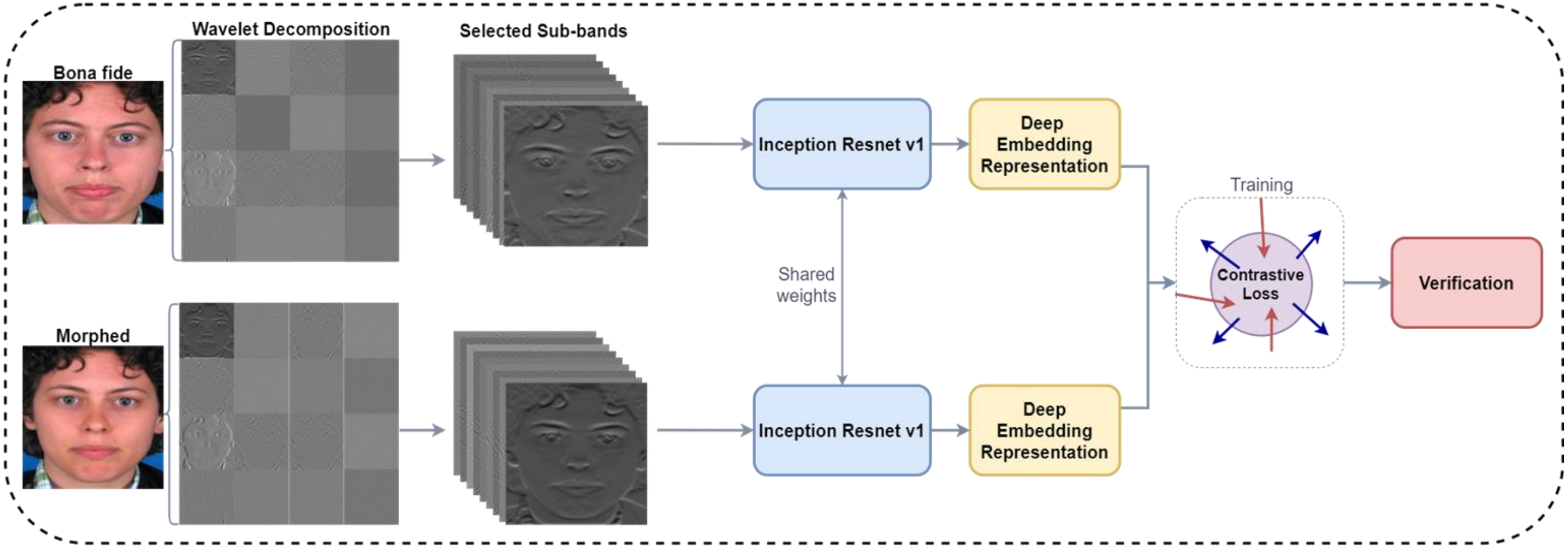
Adversarial Perturbed Morph Image

K. O'Haire, S. Soleymani, B. Chaudhary, P. Aghdaie, J. Dawson, N. Nasrabadi, "Adversarially Perturbed Wavelet-based Morphed Face Generation," IEEE International Conference on Automatic Face and Gesture Recognition (FG'2021), Dec. 15-18, 2021, Jodhpur, India

WVU single and differential morph detectors



(a) single morph detection architecture

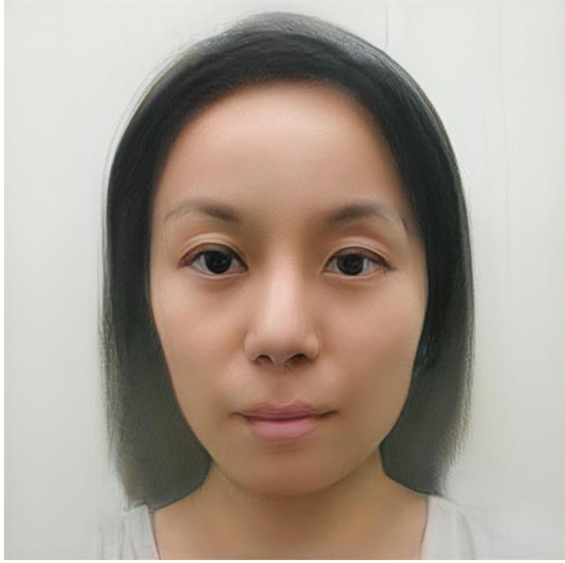


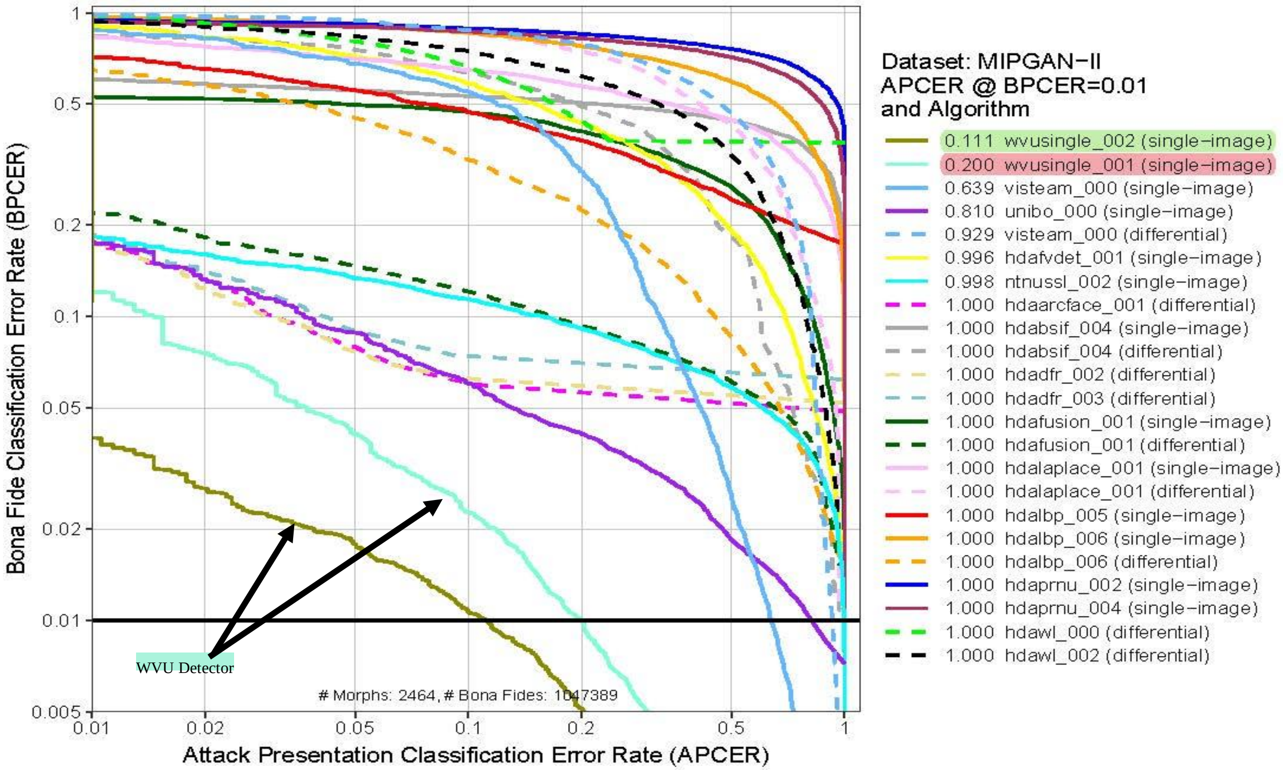
(a) differential morph detection architecture

Baaria Chaudhary, Poorya Aghdaie, Sobhan Soleymani, Jeremy Dawson, Nasser M. Nasrabadi, “Differential Morph Face Detection using Discriminative Wavelet Sub-bands,” *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshop on Biometrics (CVPRW)*,” Virtual, June 19-25, 2021.

WVUsingle_001+ WVUsingle_002.v1 NIST Performance -MIPGAN-II

APCER – percentage of morphs mislabeled as bona fide
BPCER – percentage of bona fides mis labeled as morph




Dataset: MIPGAN−II
APCER @ BPCER=0.01
and Algorithm
0.111 wvusingle_002 (single−image)
0.200 wvusingle_001 (single−image)
0.639 visteam_000 (single−image)
0.810 unibo_000 (single−image)
0.929 visteam_000 (differential)
0.996 hdafvdet_001 (single−image)
0.998 ntnussl_002 (single−image)
1.000 hdaarcface_001 (differential)
1.000 hdabsif_004 (single−image)
1.000 hdabsif_004 (differential)
1.000 hdadfr_002 (differential)
1.000 hdadfr_003 (differential)
1.000 hdafusion_001 (single−image)
1.000 hdafusion_001 (differential)
1.000 hdalaplace_001 (single−image)
1.000 hdalaplace_001 (differential)
1.000 hdalbp_005 (single−image)
1.000 hdalbp_006 (single−image)
1.000 hdalbp_006 (differential)
1.000 hdaprnu_002 (single−image)
1.000 hdaprnu_004 (single−image)
1.000 hdawl_000 (differential)
1.000 hdawl_002 (differential)
Bona Fide Classification Error Rate (BPCER)
1
0.5
0.2
0.1
0.05
0.02
0.01
0.005
0.01
0.02
0.05
0.1
0.2
0.5
1
Attack Presentation Classification Error Rate (APCER)
# Morphs: 2464, # Bona Fides: 1047389
WVU Detector

DET plot. This charts plots BPCER as a function of APCER. The x-axis is the rate morphs are not detected and the y-axis is the rate that bona fide images are falsely classified as morphs. The horizontal black line represents BPCER=0.01.

Publications

1. Samuel Price, Sobhan Soleymani, Nasser M. Nasrabadi, " Landmark Enforcement and Style Manipulation for Generative Morphing," *IEEE Int. Joint Conference on Biometrics (IJCB'22)*, Oct. 10-13, 2022.
2. Hossein Kashiani, Shoaib Meraj Sami, Sobhan Soleymani, Nasser M. Nasrabadi, "Robust Ensemble Morph Detection with Domain Generalization," *IEEE Int. Joint Conference on Biometrics (IJCB'22)*, Oct. 10-13, 2022.
3. Poorya Aghdaie, Sobhan Soleymani, Baaria Chaudhary, Jeremy Dawson, Nasser M. Nasrabadi, “Morph Detection Enhanced by Structured Group Sparsity,” *IEEE Winter Conference on Applications of Computer Vision (WACV 2022)*, workshop WACVMAPA2022, Jan. 4-8, 2022, Waikoloa, Hawaii.
4. Kelsey O'Haire, Sobhan Soleymani, Baaria Chaudhary, Poorya Aghdaie, Jeremy Dawson, Nasser M Nasrabadi, “Adversarially Perturbed Wavelet-based Morphed Face Generation,” IEEE International Conference on Automatic Face and Gesture Recognition (FG’2021), Dec. 15-18, 2021.Poorya Aghdaie, Baaria Chaudhary, Sobhan Soleymani, Jeremy Dawson, Nasser M. Nasrabadi, "Attention Aware Wavelet-based Detection of Morphed Face Images," *IEEE Int. Joint Conference on Biometrics (IJCB'21)*, Aug. 4-7, 2021.
5. Baaria Chaudhary, Poorya Aghdaie, Sobhan Soleymani, Jeremy Dawson, Nasser M. Nasrabadi, “Differential Morph Face Detection using Discriminative Wavelet Sub-bands,” *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshop on Biometrics (CVPRW)*,” Virtual, June 19-25, 2021.
6. Poorya Aghdaie, Baaria Chaudhary, Sobhan Soleymani, Jeremy Dawson, Nasser M. Nasrabadi, “Detection of Morphed Face Images Using Discriminative Wavelet Sub-bands,” the *9th IEEE International Workshop on Biometrics and Forensics (IWBF21)*, May 6-7, 2021..
7. Sobhan Soleymani, Baaria Chaudhary, Ali Dabouei, Jeremy Dawson, Nasser M. Nasrabadi, “Differential Morphed Face Detection Using Deep Siamese Networks,” *ICPR Workshop, MultiMedia FORensics in the WILD (MMForWILD’2020)*, Jan. 10-15, 2021.
8. Sobhan Soleymani, Ali Dabouei, Fariborz Taherkhani, Jeremy Dawson and Nasser M. Nasrabadi, “Mutual Information Maximization on Disentangled Representations for Differential Morph Detection,” *IEEE Winter Conference on Applications of Computer Vision (WACV 2021),* March 5-9, 2021, Waikoloa, Hawaii.

Team CU
Chen Liu (PI)
Zander Blasingame (Student)

Number of morphs CU team generated

Team	Dataset	Method	Number of Morphs
Clarkson	FRLL	Print & Scan	1,222
	MBD	StyleGAN2 Projector	2,320
	MBD	StyleGAN2 Projector, Print & Scan	932
	MBD	StyleGAN2 e4e	10,295
	Total	**Combined**	**14,760**

Accomplishment – Team Clarkson

Dataset for Constructing New Morphs

- Face dataset from Multimodal Biometric Dataset Collection (MBDC)
- 263 unique identities, 3033 total images
- Used two methods for finding latent representations
    - StyleGAN2 Projector. Uses backprop (1000 epochs) to find the closest matching latent code
    - encoder4editing (e4e). Additional network to perform encoding, trained on FFHQ
- Aligned and cropped images before passing to latent space encoders

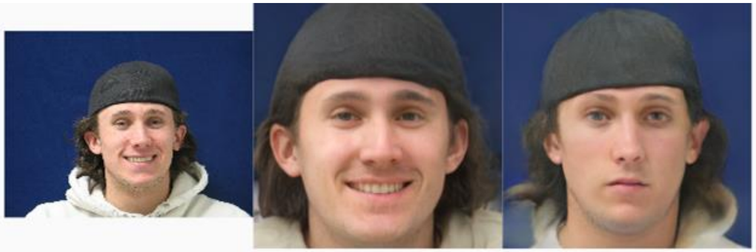


Figure 1: Real (left) vs e4e (middle) vs reconstructed (right) StyleGAN2 faces

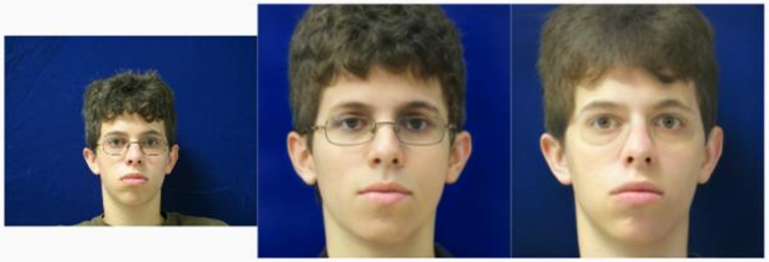
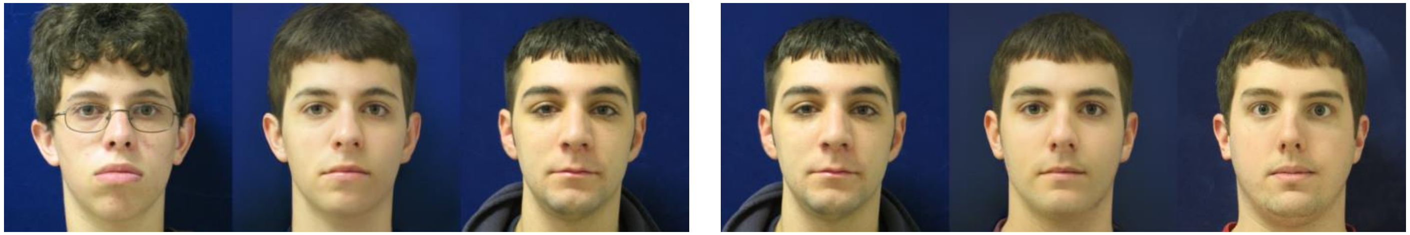


Figure 2: Real (left) vs e4e (middle) vs reconstructed (right) StyleGAN2 faces

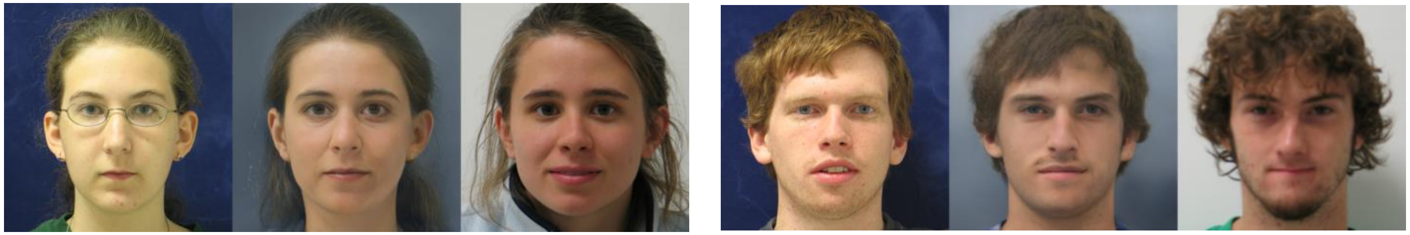
Accomplishment – Team Clarkson

Morph Generation

- StyleGAN2 based morphs using Multimodal Biometric Dataset Collection (MBDC)
- 2,320 StyleGAN2 projector-based morphs, 600 x 600 resolution
- 10,295 StyleGAN2 e4e-based morphs, high quality 1024 x 1024 resolution
- StyleGAN2 e4e morphs where chosen from top 0.2 % of morph candidates



StyleGAN2 e4e-based morphs



StyleGAN2 projector-based morphs

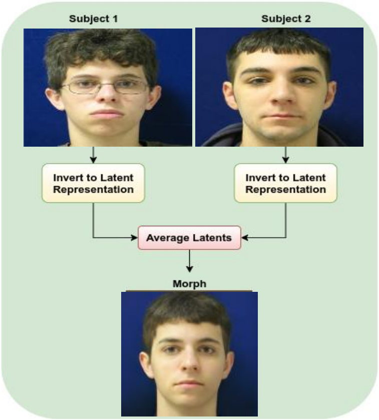
Accomplishment – Team Clarkson

FaceNet Verifier

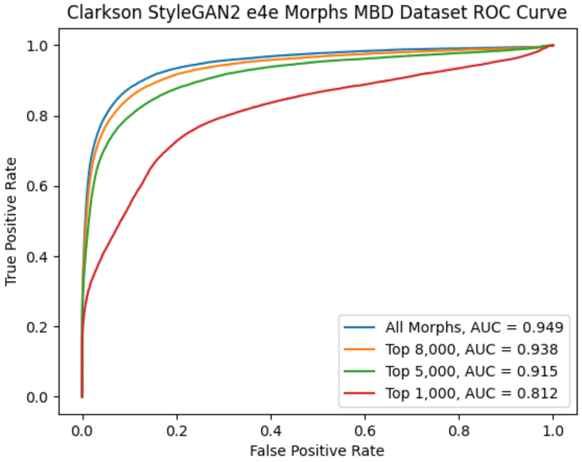
- FaceNet verifier is used to measure performance StyleGAN2 e4e Morphs

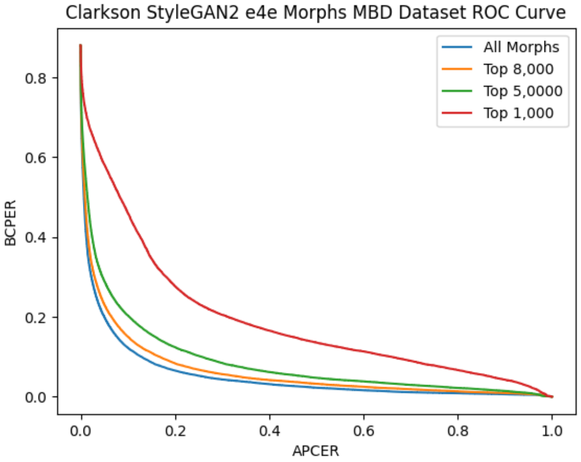
Name	Number of Morphs	AUC	APCER @ BPCER = 1%	APCER @ BPCER = 5%	APCER @ BPCER = 10%	MMPMR @ FAR = 0.1 %
All Morphs	10,295	94.9%	74.80%	26.24%	12.89%	70.67%
Top8K	8,000	93.8%	88.26%	33.81%	16.75%	85.34%
Top5K	5,000	91.5%	94.41%	48.12%	25.68%	93.40%
Top1K	1,000	81.2%	97.89%	87.11%	65.88%	98.30%

StyleGAN2 Morphing Method



Subject 1
Subject 2
Invert to Latent Representation
Invert to Latent Representation
Average Latents
Morph

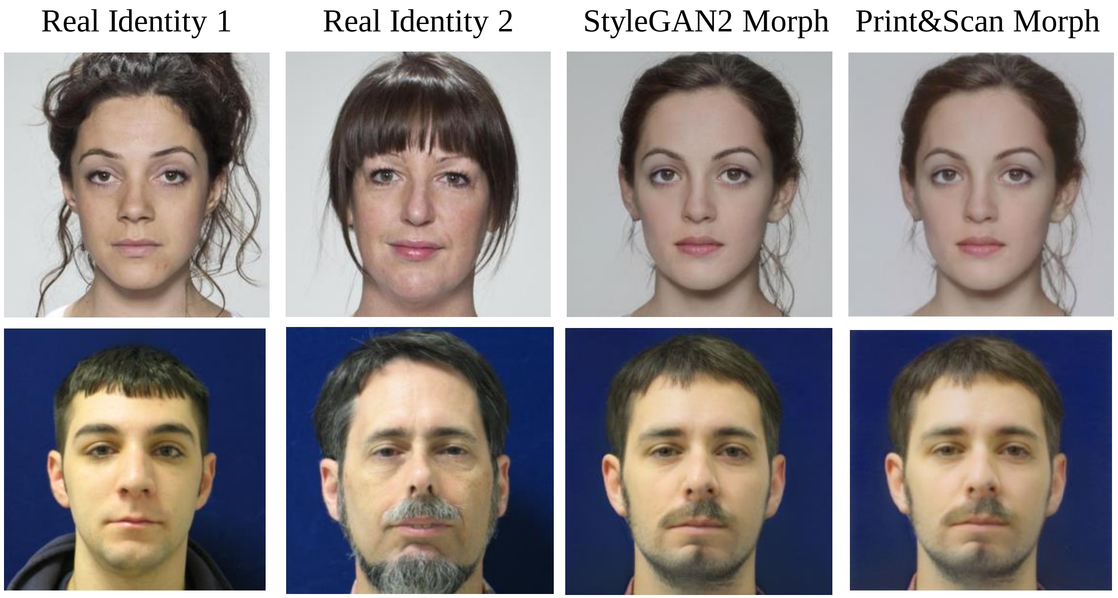




Accomplishment – Team Clarkson

Printed and Scanned Morphs

- Face datasets from Face Research Lab London (FRLL)(Top), Multimodal Biometric Dataset Collection (MBDC)(Middle).
- Printed and Scanned morphs were generated by printing StyleGAN2 morphs at 600 x 600 pixels.
  - ○ Printed with Canon Pixma Pro 100 on Canon Photo Paper Plus II
  - ○ Scanned with Epson Perfection V850 Pro



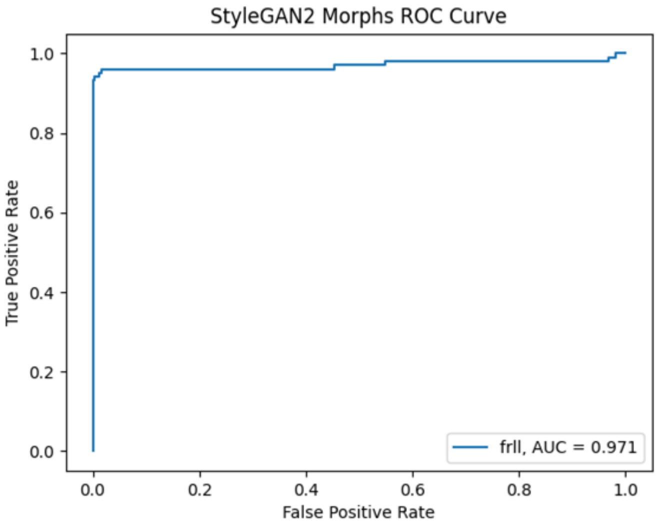
Accomplishment – Team Clarkson

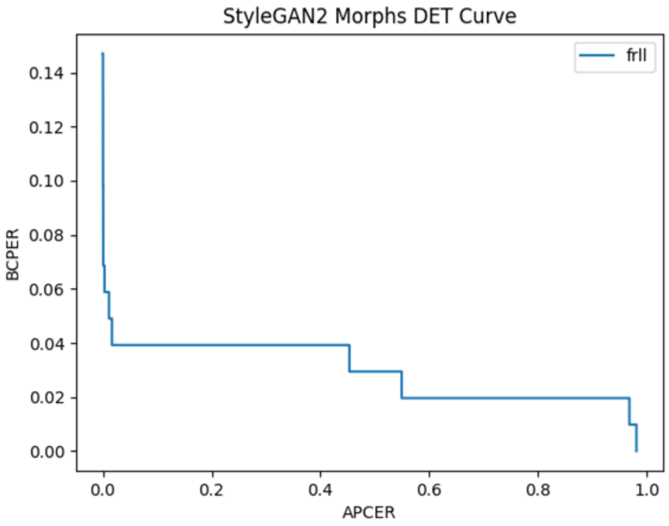
FaceNet Verifier

- FaceNet verifier is used to measure performance StyleGAN2 Print & Scan Morphs

Name	Number of Morphs	AUC	APCER @ BPCER = 1%	APCER @ BPCER = 5%	APCER @ BPCER = 10%
FRLL	1,222	97.1%	96.76%	1.13%	0.04%

- Used digital bonafide images for verification testing which leads to bias





Accomplishment – Team Clarkson

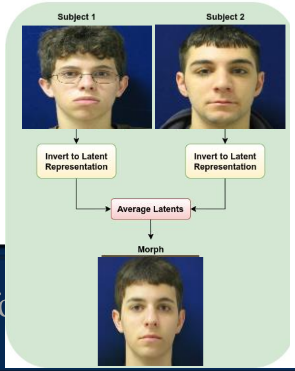
FaceNet Verifier

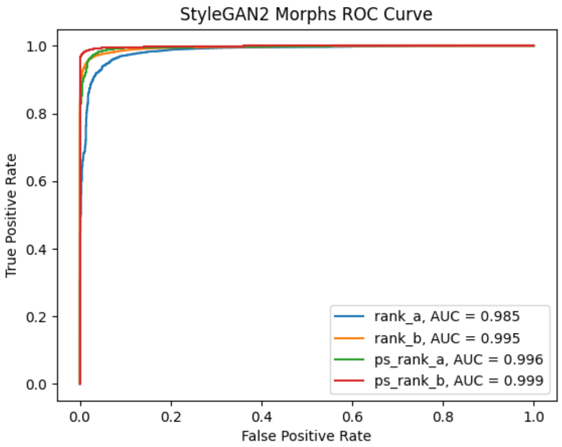
- FaceNet verifier is used to measure performance StyleGAN2 Print & Scan Morphs

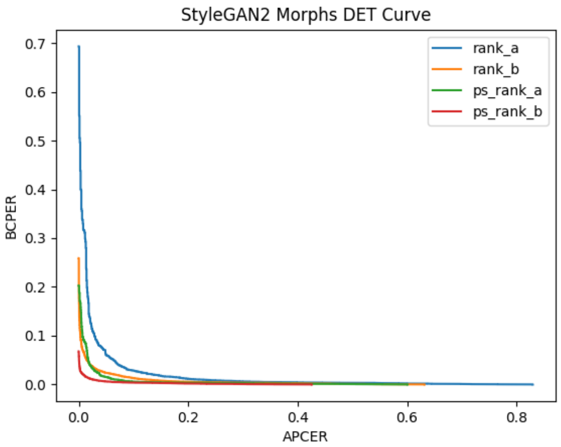
Name	Method	Number of Morphs	AUC	APCER @ BPCER = 1%	APCER @ BPCER = 5%	APCER @ BPCER = 10%
Rank B	StyleGAN2 Projector	744	99.5%	12.24%	1.46%	0.28%
Rank A	StyleGAN2 Projector	188	98.5%	22.00%	6.41%	2.90%
PS Rank B	StyleGAN2 Projector, Print & Scan	744	99.9%	2.66%	0.06%	N/A
PS Rank A	StyleGAN2 Projector, Print & Scan	188	99.6%	6.68%	1.72%	0.67%

- Used digital bonafide images for verification testing which favors digital morphs

StyleGAN2 Morphing Method


Subject 1
Subject 2
Invert to Latent Representation
Invert to Latent Representation
Average Latents
Morph


StyleGAN2 Morphs ROC Curve

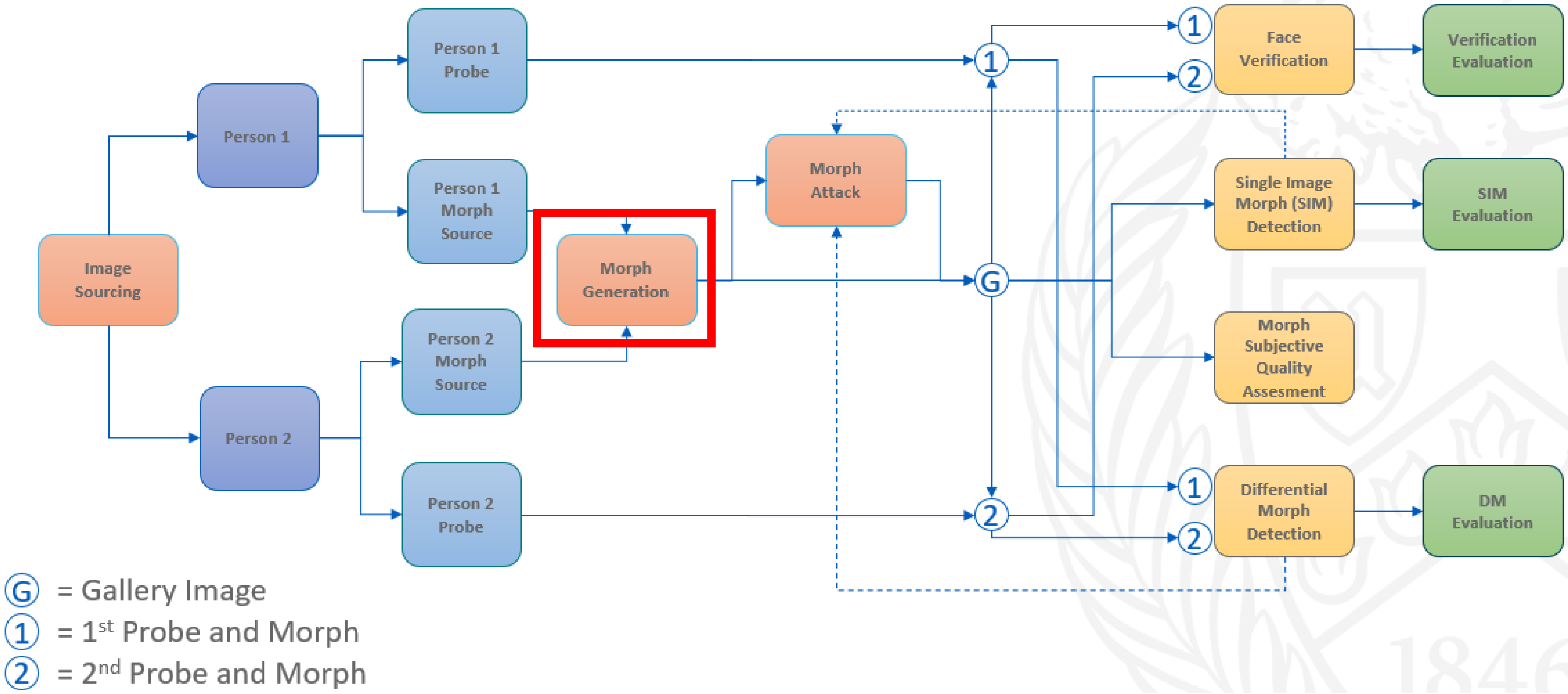

StyleGAN2 Morphs DET Curve

Publications and Products

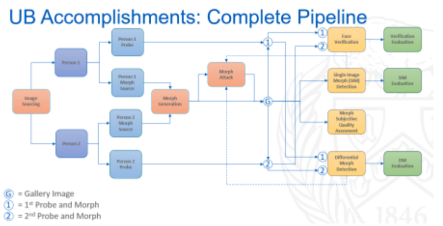
- Zander Blasingame and Chen Liu, "*Leveraging Adversarial Learning for the Detection of Morphing Attacks*," 2021 IEEE International Joint Conference on Biometrics (IJCB), 2021, pp. 1-8, doi: 10.1109/IJCB52358.2021.9484383.
- Morph Detector
  - Single morph detector (Team Clarkson): https://github.com/CamelClarkson/Morphed-Face-Detector
- Next Steps
  - Diffusion-based Morphs
    - Generated morphs using Diffusion models instead of GAN
    - Embed images into stochastic and semantic latent spaces for morphing
  - Print-Scan Morphs
    - Extending the print and scan to bonafide images for more accurate verification and testing
    - Can explore different verification algorithms

Team UB
David Doermann (PI)

UB Accomplishments: Complete Pipeline

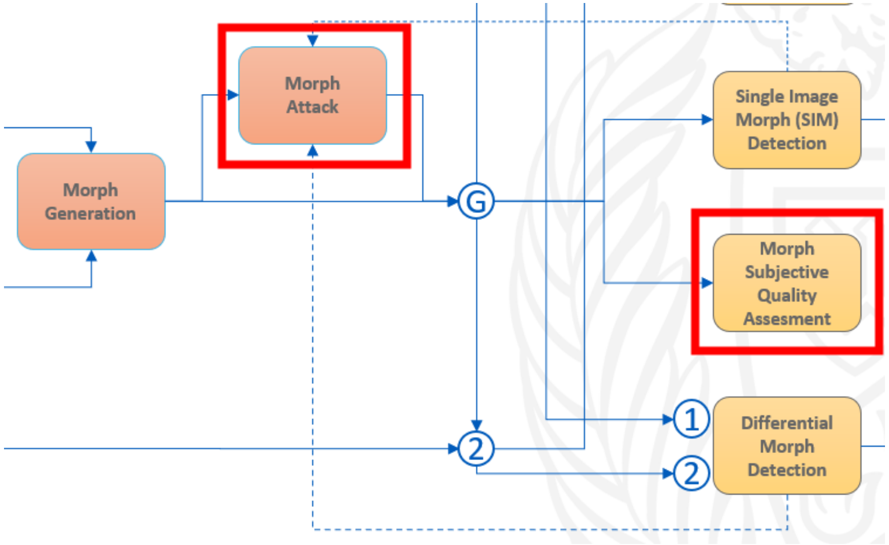


Research Focus

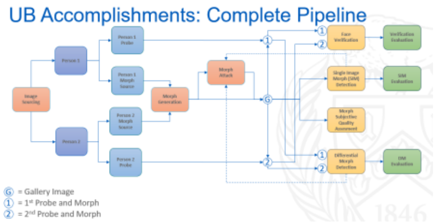


- High-Quality Morphs
  - High Visual Quality
  - Morph can pass face recognition for both subjects
  - Morph can not be detected by morph detectors

- Morph Attacks
  - Removing Trace
- Image Quality
  - Image Quality for Portraits


Morph Attack
Morph Generation
G
Single Image Morph (SIM) Detection
Morph Subjective Quality Assesment
1
2
2
Differential Morph Detection

High Quality Morph Details

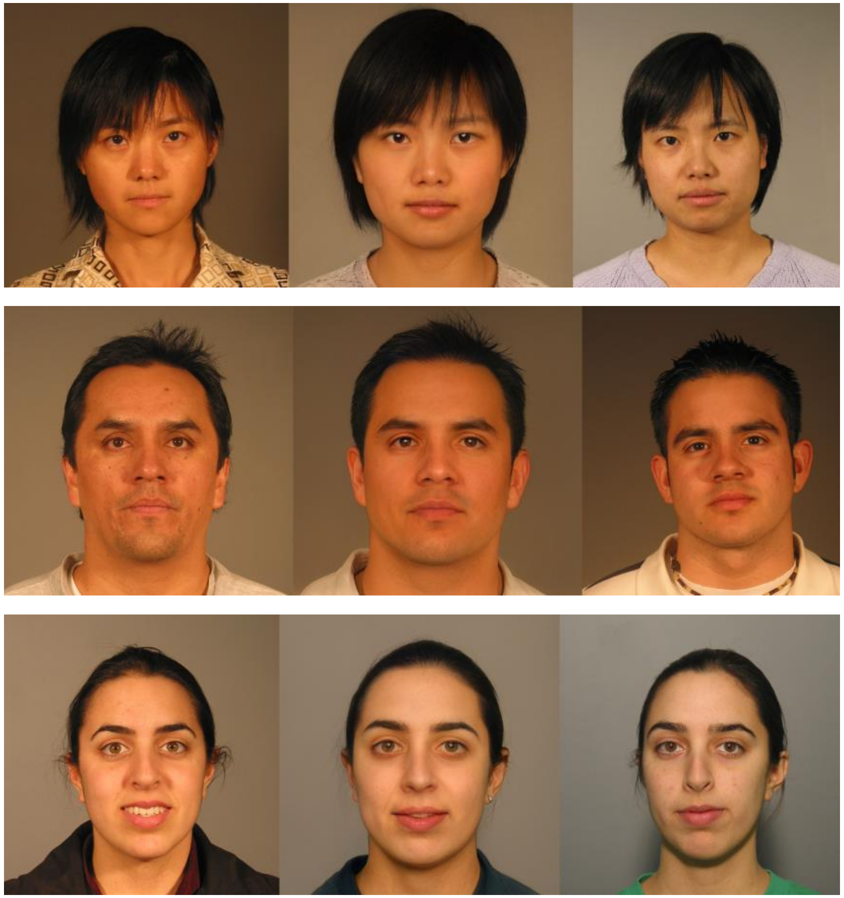


- Select 3 face images per subject based on quality and “passport” constraints
- Rank subject pairs based on similarity using ArcFace parameters
- Generate Morphs (for example w/ StyleGAN2 encoding):
  - For each pair
    - Generate with a variety of alpha ([0.4, 0.6] with step of 0.01)
    - Calculate the distance d between morphed face with a new instance of subject
    - Choose alpha based on the min absolute different of d1 and d2
    - Calculate the the average distance with the best alpha for each pair (called as *mean*)
    - Sort *mean* value and get the corresponding pairs
    - Choose top N pairs for min *mean* values
- Verify Face and Morph similarity

Automation

- Similarity Selection
- Morph Generation
- Face Verification (source vs morph)
- Morph Detection

- Generated 1000 realistic morphs
- Manual Touchup



Performance on UB FRGC StyleGAN2 morphs

Differential Morph attack detection[1]:

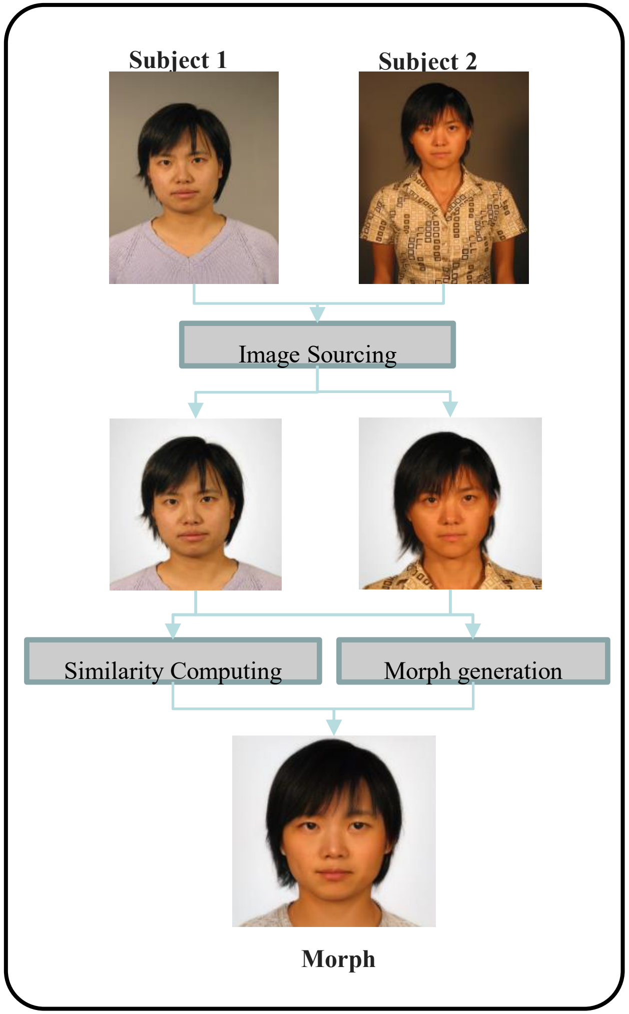
- ArcFace for face representation extraction
- SVM-based decision

Good performance for in-domain dataset (left)

- BPCER at APCER = 0.01: 0.51%
- BPCER at APCER = 0.10: 0.00%

Bad performance for cross-domain datasets (right)

- BPCER at APCER = 0.01: 60.57%
- BPCER at APCER = 0.10: 40.50%



Subject 1

Subject 2

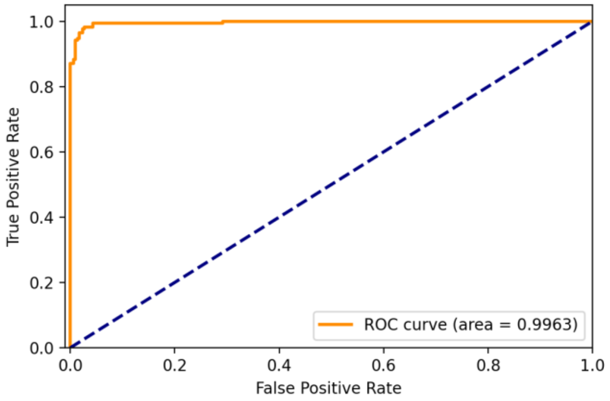
Image Sourcing

Similarity Computing

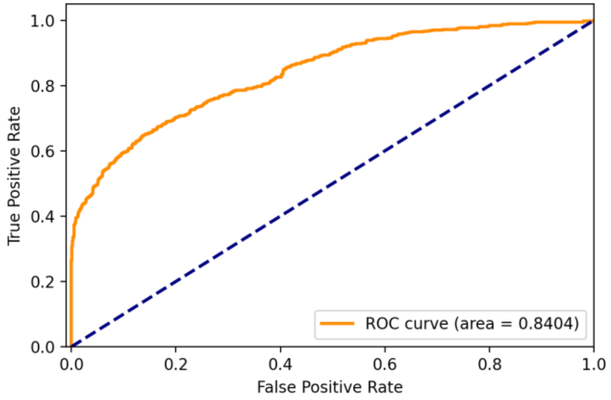
Morph generation

Morph

Training and testing on FRGC and our morphs



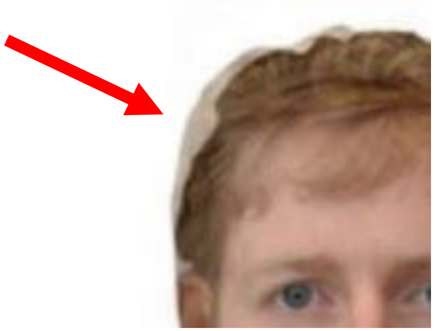
Training on Color FERET and FERET-Morphs, Testing on FRGC and our morphs



[1] Scherhag, U., Rathgeb, C., Merkle, J., & Busch, C. (2020). *IEEE Transactions on Information Forensics and Security*, *15*, 3625-3639.

Image Quality

- Need to analyze human subjective qualities of morphs as well
- Metrics are set up for Scenes or Faces, but not both



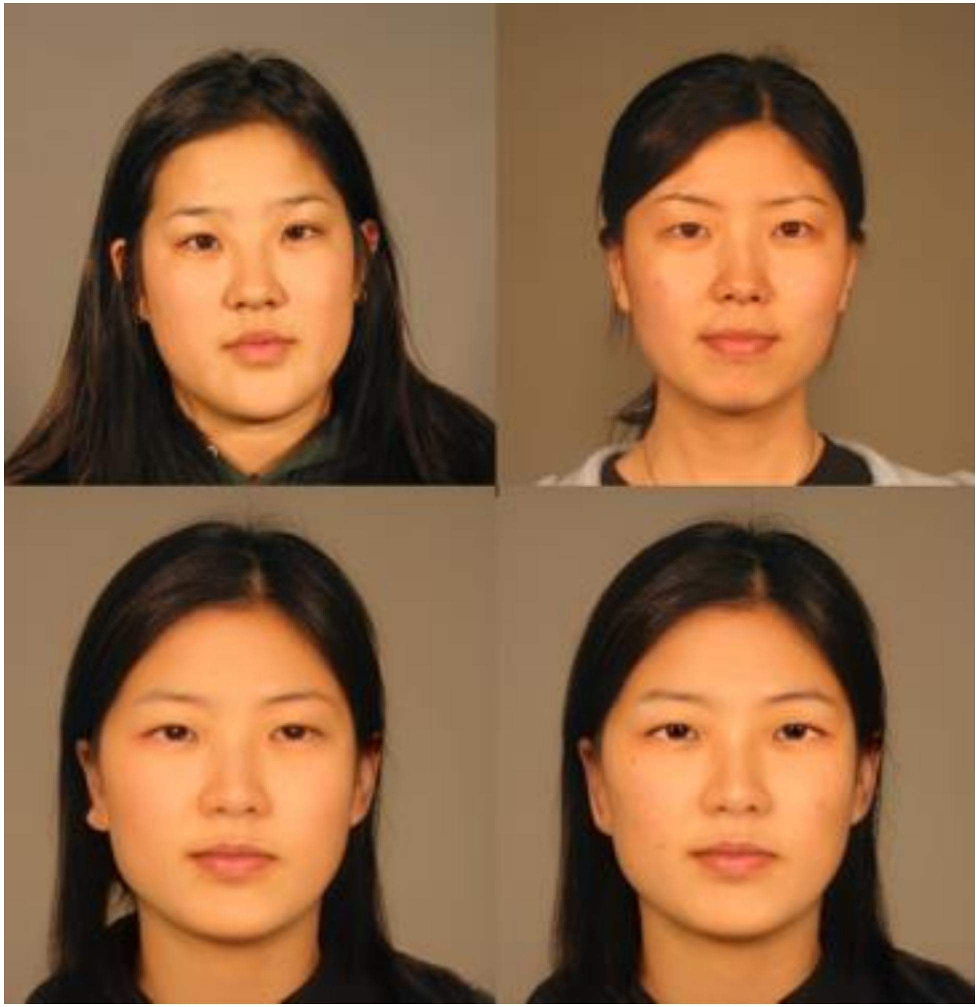
[1]

- General: Brisque,
- Face Specific: FaceQnet, FIQ, SER-FIQ

[1] Kramer, R.S.S., Mireku, M.O., Flack, T.R. *et al.* Face morphing attacks: Investigating detection with humans and computers. *Cogn. Research* **4**, 28 (2019).

Accomplishments

UB Morphs Manual Touch-up



- Objectives
  - Fix morphing artifacts
  - Introduce subtle human-observable features into morphs

Manual edit notes

- used liquify tool to change pupil shape in right eye
- used dodge tool to lighten whites of eyes and pupil reflection
- darkened edges of nose creases and bridge
- brought down nose a little
- used eyebrows from reference photos and blended them to help natural eyebrow shape and look
- used lasso tool to trim small section of the bottom of ear to match other
- copied the right ear and flipped it horizontally to place on left, then blend
- added acne/scarring/dark spots/freckles from reference images

Questions?