IEEE International Joint Conference on Biometrics

Diffusion Morphs (DiM)

Leveraging Diffusion for Strong and High-Quality Face Morphing Attacks

---

Zander W. Blasingame   Chen Liu

Clarkson University
Potsdam, NY, USA

09.18.2024

Introduction

Face Morphing

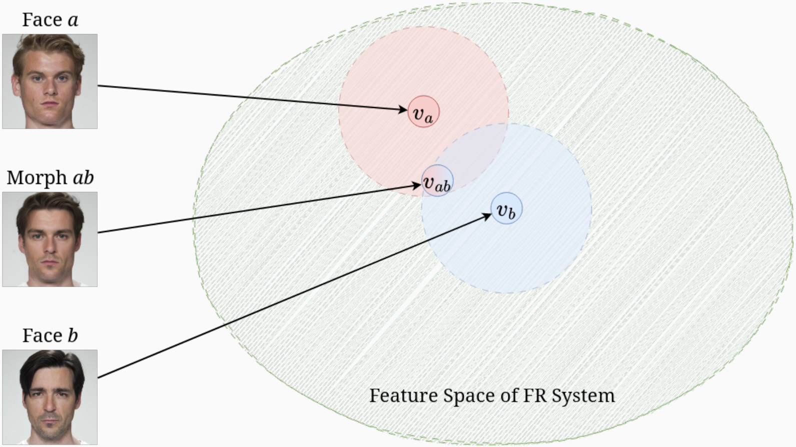


Figure 1: Images from FRLL[1] dataset. Morph generated via DiM.

[1] Lisa DeBruine and Benedict Jones. "Face Research Lab London Set". In: (May 2017). DOI: 10.6084/m9.figshare.5047666.v5. URL: https://figshare.com/articles/dataset/Face_Research_Lab_London_Set/5047666.

Morph Creation Pipeline

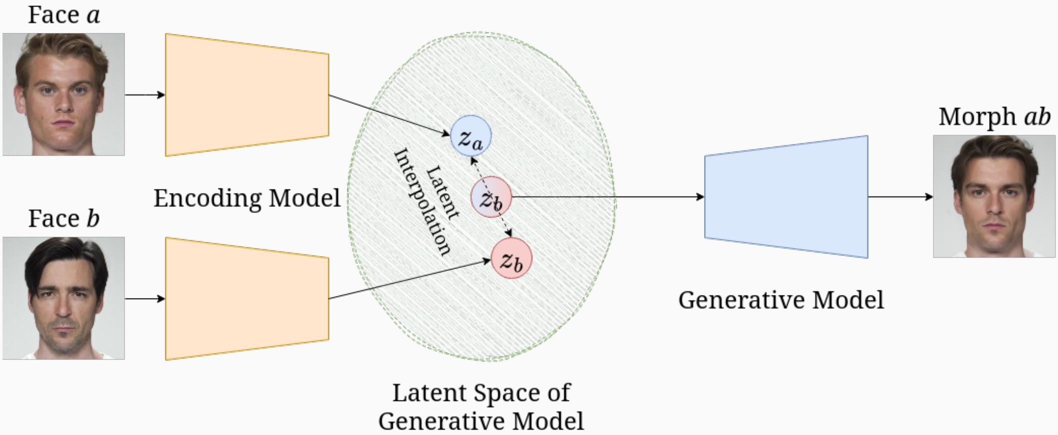


Figure 2: General morph creation pipeline using generative models.

Diffusion Process

- Forward diffusion process is governed by the Itô SDE

$$\mathrm{d}\mathbf{x}_t = f(t)\mathbf{x}_t \,\mathrm{d}t + g(t)\,\mathrm{d}\mathbf{w}_t \tag{1}$$

where $\{\mathbf{w}_t\}_{t\in[0,T]}$ is the standard Wiener process on $[0, T]$

[2] Yang Song et al. "Score-Based Generative Modeling through Stochastic Differential Equations". In: *International Conference on Learning Representations*. 2021. URL: https://openreview.net/forum?id=PxTIG12RRHS.

[3] Tim Salimans and Jonathan Ho. "Progressive Distillation for Fast Sampling of Diffusion Models". In: *International Conference on Learning Representations*. 2022. URL: https://openreview.net/forum?id=TIdIXIpzhoI.

Diffusion Process

- The diffusion equation can be reversed with

$$\mathrm{d}\mathbf{x}_t = \left[f(t)\mathbf{x}_t - g^2(t)\nabla_{\mathbf{x}} \log p_t(\mathbf{x}_t)\right] \mathrm{d}t + g(t)\, \mathrm{d}\check{\mathbf{w}}_t \tag{2}$$

  where $\check{\mathbf{w}}_t$ is the *backwards* Wiener process defined as $\check{\mathbf{w}}_t := \mathbf{w}_t - \mathbf{w}_T$
- The marginal distributions $p_t(\mathbf{x})$ follow an associated ODE known as the *probability flow* ODE[2]

$$\frac{\mathrm{d}\mathbf{x}_t}{\mathrm{d}t} = f(t)\mathbf{x}_t - \frac{1}{2}g^2(t)\nabla_{\mathbf{x}} \log p_t(\mathbf{x}_t) \tag{3}$$

[2] Yang Song et al. "Score-Based Generative Modeling through Stochastic Differential Equations". In: *International Conference on Learning Representations*. 2021. URL: https://openreview.net/forum?id=PxTIG12RRHS.

[3] Tim Salimans and Jonathan Ho. "Progressive Distillation for Fast Sampling of Diffusion Models". In: *International Conference on Learning Representations*. 2022. URL: https://openreview.net/forum?id=TIdIXIpzhoI.

Diffusion Process

- Often the Variance Preserving (VP) framework is used where the drift and diffusion coefficients are

$$f(t) = \frac{\mathrm{d}\log\alpha_t}{\mathrm{d}t} \tag{4}$$

$$g^2(t) = \frac{\mathrm{d}\sigma_t^2}{\mathrm{d}t} - 2\frac{\mathrm{d}\log\alpha_t}{\mathrm{d}t}\sigma_t^2 \tag{5}$$

  for some noise schedule α_t, σ_t
- Sampling the forward trajectory then simplifies to

$$\mathbf{x}_t = \alpha_t\mathbf{x}_0 + \sigma_t\boldsymbol{\epsilon}_t \qquad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \tag{6}$$

[2] Yang Song et al. "Score-Based Generative Modeling through Stochastic Differential Equations". In: *International Conference on Learning Representations*. 2021. URL: https://openreview.net/forum?id=PxTIG12RRHS.

[3] Tim Salimans and Jonathan Ho. "Progressive Distillation for Fast Sampling of Diffusion Models". In: *International Conference on Learning Representations*. 2022. URL: https://openreview.net/forum?id=TIdIXIpzhoI.

Diffusion Process

- Learning the score $\nabla_{\mathbf{x}} \log p_t(\mathbf{x}_t)$ is similar to learning the noise $\boldsymbol{\epsilon}$

$$\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) \approx -\sigma_t \nabla_{\mathbf{x}} \log p_t(\mathbf{x}_t) \tag{7}$$

  or some other closely related quantity like $\mathbf{x}_0$-prediction[3]
- Train a U-Net, $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$, to learn the added noise

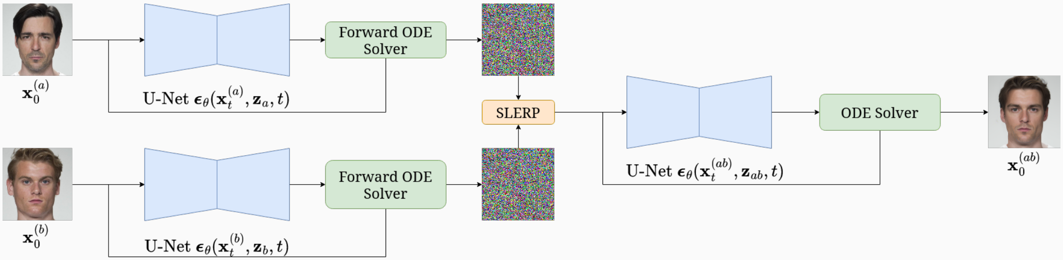
$$\hat{\theta} = \arg\min_{\theta} \quad \mathbb{E}_{\substack{\mathbf{x}_0 \sim p(\mathbf{x}_0) \\ \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})}} \left[\|\boldsymbol{\epsilon}_t - \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\|_2^2 \right] \tag{8}$$

[2] Yang Song et al. "Score-Based Generative Modeling through Stochastic Differential Equations". In: *International Conference on Learning Representations*. 2021. URL: https://openreview.net/forum?id=PxTIG12RRHS.

[3] Tim Salimans and Jonathan Ho. "Progressive Distillation for Fast Sampling of Diffusion Models". In: *International Conference on Learning Representations*. 2022. URL: https://openreview.net/forum?id=TIdIXIpzhoI.

Diffusion Morphs (DiM)

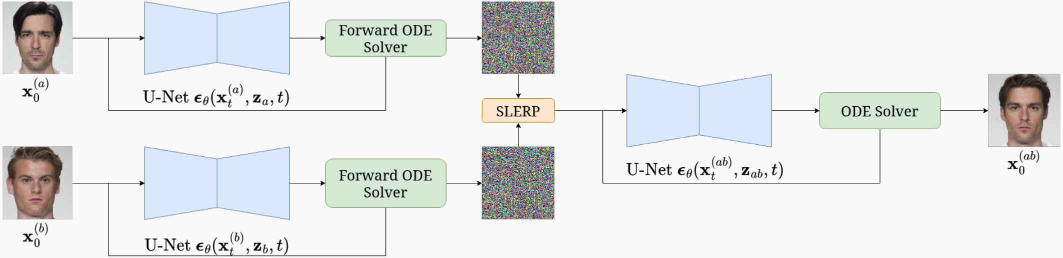
Face Morphing with Diffusion



- Encode bona fide images

$$\mathbf{z}_{\{a,b\}} = E(\mathbf{x}_0^{(\{a,b\})}) \tag{9}$$

Face Morphing with Diffusion

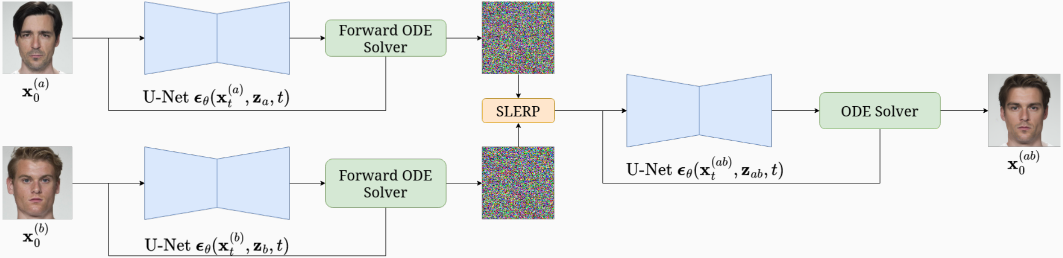


- Let $\Phi(\mathbf{x}_0, \mathbf{z}, \mathbf{h}_\theta, \{t_n\}_{n=1}^N) \mapsto \mathbf{x}_T$ denote a numerical ODE solver with
    1. Initial image $\mathbf{x}_0$
    2. Latent representation of $\mathbf{x}_0$, $\mathbf{z} = E(\mathbf{x}_0)$
    3. Denoising U-Net conditioned on $\mathbf{z}$, $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{z}, t)$
    4. The PF ODE given by

$$\mathbf{h}_\theta(\mathbf{x}_t, \mathbf{z}, t) = f(t)\mathbf{x}_t + \frac{g^2(t)}{2\sigma_t}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{z}, t) \tag{10}$$

    5. N timesteps $\{t_n\}_{n=1}^N \subseteq [0, T]$

Face Morphing with Diffusion

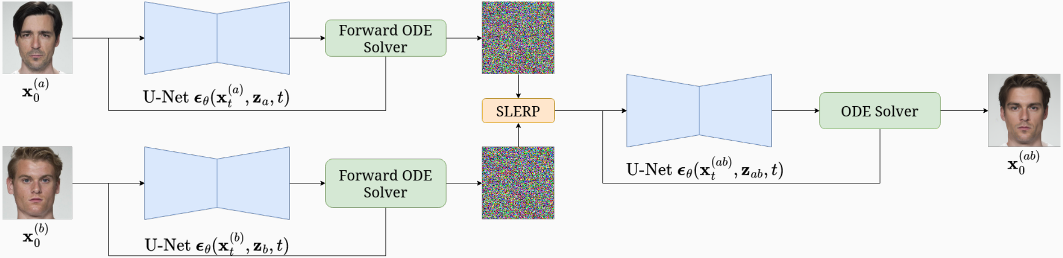


- Encode images solving the PF ODE as time runs *forwards*

$$\mathbf{x}_T^{(\{a,b\})} = \Phi(\mathbf{x}_0^{(\{a,b\})}, \mathbf{z}_{\{a,b\}}, \mathbf{h}_\theta, \{t_n\}_{n=1}^{N_F}) \quad (11)$$

with N_F encoding steps and $t_n < t_{n+1}$

Face Morphing with Diffusion



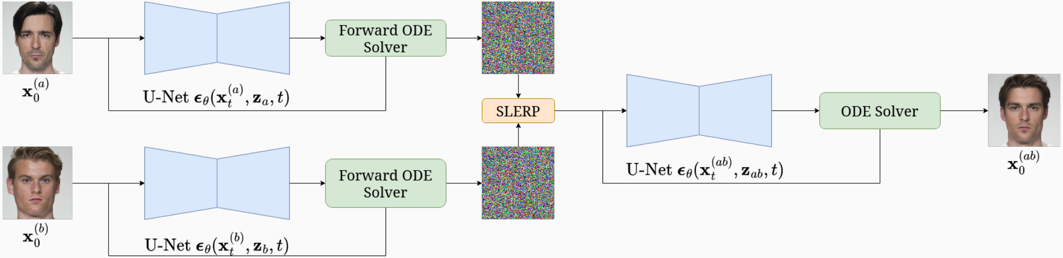
- Morph the latent representations

$$\mathbf{x}_T^{(ab)} = \mathrm{slerp}(\mathbf{x}_T^{(a)}, \mathbf{x}_T^{(b)}; \gamma) \tag{12}$$

$$\mathbf{z}_{ab} = \mathrm{lerp}(\mathbf{z}_a, \mathbf{z}_b; \gamma) \tag{13}$$

by a factor of $\gamma = 0.5$

Face Morphing with Diffusion

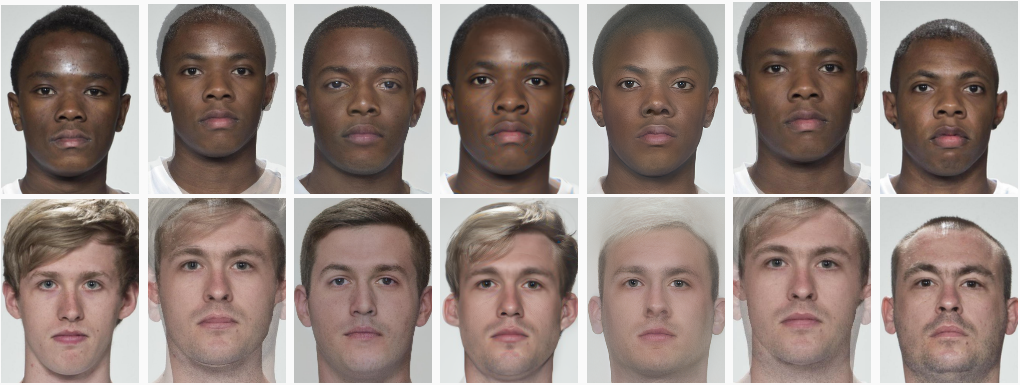


- Create morph by solving the PF ODE as time runs *backwards*

$$\mathbf{x}_0^{(ab)} = \Phi(\mathbf{x}_T^{(ab)}, \mathbf{z}_{ab}, \mathbf{h}_\theta, \{\tilde{t}_n\}_{n=1}^N) \tag{14}$$

with N sampling steps and $\tilde{t}_n > \tilde{t}_{n+1}$

Visual Comparison to Other Morphing Attacks



(a) Identity a (b) OpenCV (c) StyleGAN2 (d) DiM (e) MIPGAN-II (f) FaceMorpher (g) Identity b

Figure 3: Comparison across different morphing algorithms of two identity pairs from the FRLL dataset.

Quantitative Comparison

Table 1: Vulnerability of different FR systems across different morphing attacks on the SYN-MAD 2022 dataset. FMR $= 0.1\%$.

	MMPMR ($\uparrow$)		
Morphing Attack	**AdaFace** [8]	**ArcFace** [6]	**ElasticFace** [4]
FaceMorpher [7]	89.78	87.73	89.57
OpenCV [7]	94.48	92.43	94.27
MIPGAN-I [13]	72.19	77.51	66.46
MIPGAN-II [13]	70.55	72.19	65.24
DiM [3]	92.23	90.18	93.05

- Mated Morph Presentation Match Rate (MMPMR)

$$M(\delta) = \frac{1}{M} \sum_{n=1}^{M} \left\{ \left[\min_{n \in \{1,\ldots,N_m\}} S_m^n \right] > \delta \right\} \tag{15}$$

where δ is the verification threshold, S_m^n is the similarity score of the n-th subject of morph m, N_m is the total number of contributing subjects to morph m, and M is the total number of morphed images.

Ablation Study

Table 2: Ablation study on the ability to detect morphing attacks.

	Included in the Training Set					Detection Accuracy (↓)				
Dataset	**DiM**	**FaceMorpher**	**MIPGAN-II**	**OpenCV**	**StyleGAN2**	**DiM**	**FaceMorpher**	**MIPGAN-II**	**OpenCV**	**StyleGAN2**
FERET [9]	✗	✓	✓	✓	✓	**72.73**	99.23	100	99.95	99.33
	✓	✗	✓	✓	✓	99.9	76.39	100	99.85	99.64
	✓	✓	✗	✓	✓	99.69	99.38	100	99.95	99.54
	✓	✓	✓	✗	✓	99.74	99.48	100	99.74	99.43
	✓	✓	✓	✓	✗	99.74	98.56	99.9	99.74	87.89
FRGC [10]	✗	✓	✓	✓	✓	**75.89**	99.98	99.97	99.9	99.93
	✓	✗	✓	✓	✓	99.95	99.48	100	99.9	99.95
	✓	✓	✗	✓	✓	99.83	99.85	99.82	99.8	99.85
	✓	✓	✓	✗	✓	99.93	100	100	99.23	99.93
	✓	✓	✓	✓	✗	99.93	99.93	99.94	99.88	97.83
FRLL [5]	✗	✓	✓	✓	✓	**13.96**	99.58	99.32	99.65	99.65
	✓	✗	✓	✓	✓	99.23	99.09	98.91	99.37	99.44
	✓	✓	✗	✓	✓	99.09	98.95	98.24	99.02	99.09
	✓	✓	✓	✗	✓	99.51	99.44	99.19	99.16	99.58
	✓	✓	✓	✓	✗	99.93	99.86	99.86	99.93	95.02

Conclusions

- DiM creates morphs with high visual fidelity
- DiM outperforms GAN-based morphs
- DiM is difficult to detect if not explicitly trained against
- Flexible generation due to iterative nature
- Slow inference speed due to multiple iterations

Related Work

Since our initial work there have been several extensions and improvements on DiM

Fast-DiM[4] High-order ODE solvers for faster sampling, reduces NFE from 350 to 150

Morph-PIPE[5] Brute force search for optimal γ w.r.t. an identity loss, increased MMPMR

Greedy-DiM[6] Greedy guided generation for DiM, 100% MMPMR on SYN-MAD 22

---

[4] Zander W. Blasingame and Chen Liu. "Fast-DiM: Towards Fast Diffusion Morphs". In: *IEEE Security & Privacy* (2024), pp. 2–13. DOI: 10.1109/MSEC.2024.3410112.

[5] Haoyu Zhang et al. "Morph-PIPE: Plugging in Identity Prior to Enhance Face Morphing Attack Based on Diffusion Model". In: *Norwegian Information Security Conference (NISK)*. 2023.

[6] Zander W. Blasingame and Chen Liu. "Greedy-DiM: Greedy Algorithms for Unreasonably Effective Face Morphs". In: *arXiv e-prints*, arXiv:2404.06025 (Apr. 2024), arXiv:2404.06025. DOI: 10.48550/arXiv.2404.06025. arXiv: 2404.06025 [cs.CV].

Questions?

Code and project page for DiM

Further reading on DiM

References

[1] Zander W. Blasingame and Chen Liu. "Fast-DiM: Towards Fast Diffusion Morphs". In: *IEEE Security & Privacy* (2024), pp. 2–13. DOI: 10.1109/MSEC.2024.3410112.

[2] Zander W. Blasingame and Chen Liu. "Greedy-DiM: Greedy Algorithms for Unreasonably Effective Face Morphs". In: *arXiv e-prints*, arXiv:2404.06025 (Apr. 2024), arXiv:2404.06025. DOI: 10.48550/arXiv.2404.06025. arXiv: 2404.06025 [cs.CV].

[3] Zander W. Blasingame and Chen Liu. "Leveraging Diffusion for Strong and High Quality Face Morphing Attacks". In: *IEEE Transactions on Biometrics, Behavior, and Identity Science* 6.1 (2024), pp. 118–131. DOI: 10.1109/TBIOM.2024.3349857.

[4] Fadi Boutros et al. "ElasticFace: Elastic Margin Loss for Deep Face Recognition". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*. June 2022, pp. 1578–1587.

[5] Lisa DeBruine and Benedict Jones. "Face Research Lab London Set". In: (May 2017). DOI: 10.6084/m9.figshare.5047666.v5. URL: https://figshare.com/articles/dataset/Face_Research_Lab_London_Set/5047666.

[6] Jiankang Deng et al. "Arcface: Additive angular margin loss for deep face recognition". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019, pp. 4690–4699.

[7] Marco Huber et al. "SYN-MAD 2022: Competition on Face Morphing Attack Detection Based on Privacy-aware Synthetic Training Data". In: *2022 IEEE International Joint Conference on Biometrics (IJCB)*. 2022, pp. 1–10. DOI: 10.1109/IJCB54206.2022.10007950.

[8] Minchul Kim, Anil K Jain, and Xiaoming Liu. "AdaFace: Quality Adaptive Margin for Face Recognition". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

[9] P. Phillips et al. "The FERET database and evaluation procedure for face-recognition algorithms". In: *Image Vis. Comput.* 16 (1998), pp. 295–306.

References iii

[10] P.J. Phillips et al. "Overview of the face recognition grand challenge". In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. Vol. 1. 2005, 947–954 vol. 1. DOI: 10.1109/CVPR.2005.268.

[11] Tim Salimans and Jonathan Ho. "Progressive Distillation for Fast Sampling of Diffusion Models". In: *International Conference on Learning Representations*. 2022. URL: https://openreview.net/forum?id=TIdIXIpzhoI.

[12] Yang Song et al. "Score-Based Generative Modeling through Stochastic Differential Equations". In: *International Conference on Learning Representations*. 2021. URL: https://openreview.net/forum?id=PxTIG12RRHS.

[13] Haoyu Zhang et al. "MIPGAN—Generating Strong and High Quality Morphing Attacks Using Identity Prior Driven GAN". In: *IEEE Transactions on Biometrics, Behavior, and Identity Science* 3.3 (2021), pp. 365–383. DOI: 10.1109/TBIOM.2021.3072349.

[14] Haoyu Zhang et al. "Morph-PIPE: Plugging in Identity Prior to Enhance Face Morphing Attack Based on Diffusion Model". In: *Norwegian Information Security Conference (NISK)*. 2023.