Diffusion Morphs (DiM)

Diffusion is all you need for highly effective face morphs

Zander W. Blasingame Stephanie Schuckers Chen Liu

Clarkson University
Potsdam, NY, USA

11.06.2024

Introduction

Face Morphing

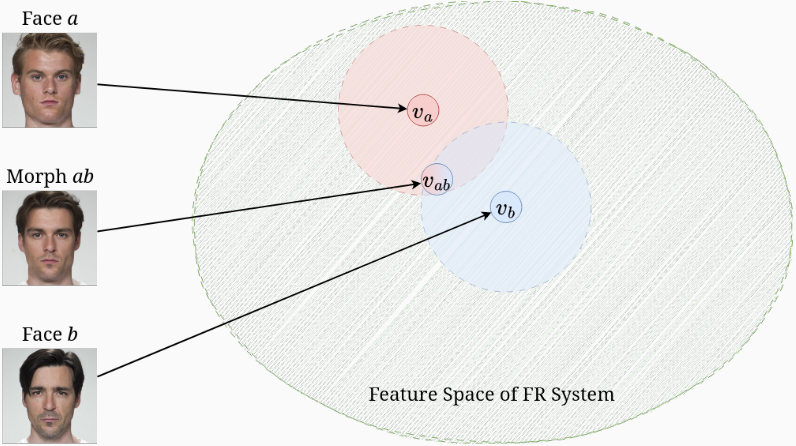


Figure 1: Images from FRLL[1] dataset. Morph generated via DiM.

[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

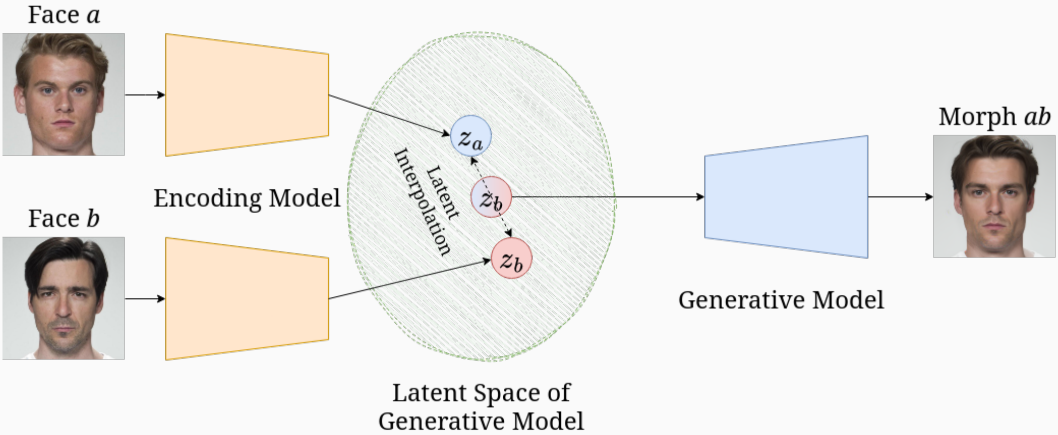
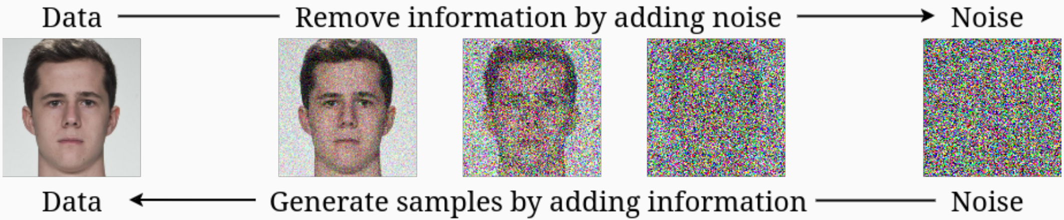


Figure 2: General morph creation pipeline using generative models.

Diffusion Models



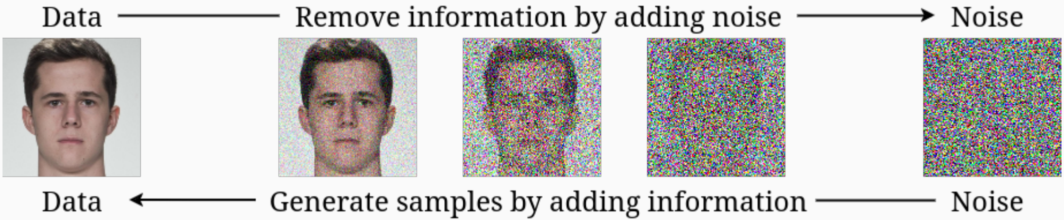
- Forward diffusion process is governed by the Itô SDE

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

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.

Diffusion Models



- The diffusion equation can be reversed with

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

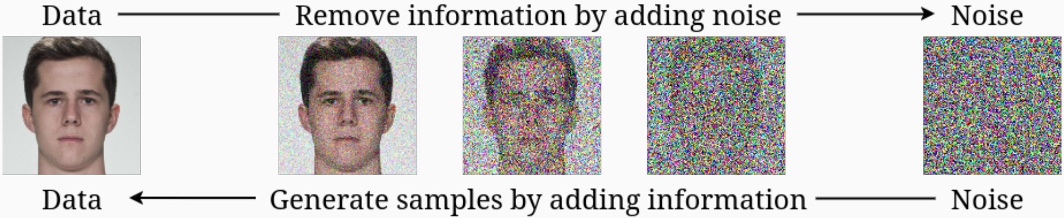
where $\bar{\mathbf{w}}_t$ is the *reverse* Wiener process and '$\mathrm{d}t$' is a *negative* timestep.

- The marginal distributions $p_t(\mathbf{x})$ follow 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.

Diffusion Models



- Train the model via score-matching to learn $\nabla_{\mathbf{x}} \log p_t(\mathbf{x}_t)$.
- This is similar to learning the noise $\boldsymbol{\epsilon}$, *i.e.*,

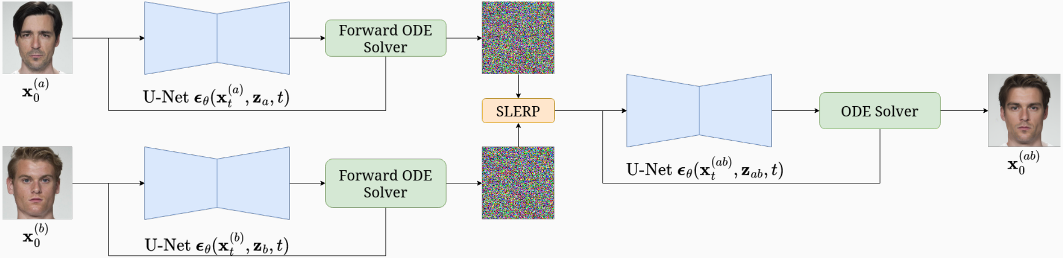
$$\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) \approx -\sigma_t \nabla_{\mathbf{x}} \log p_t(\mathbf{x}_t), \tag{4}$$

with $\mathbf{x}_t = \alpha_t \mathbf{x}_0 + \sigma_t \boldsymbol{\epsilon}$.

[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.

Diffusion Morphs (DiM)

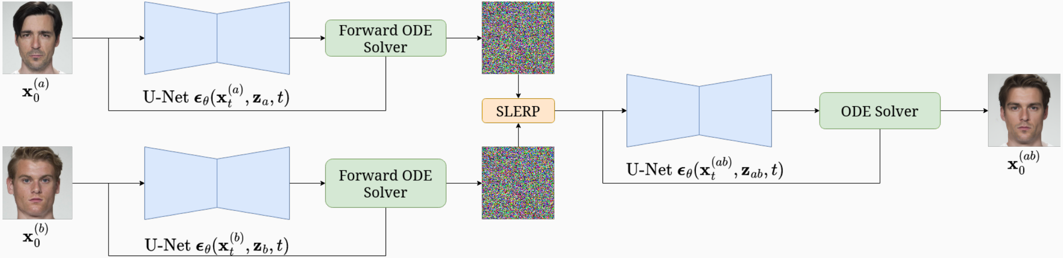
Face Morphing with Diffusion



- Encode bona fide images:

$$\mathbf{z}_{\{a,b\}} = E(\mathbf{x}_0^{(\{a,b\})}). \tag{5}$$

Face Morphing with Diffusion

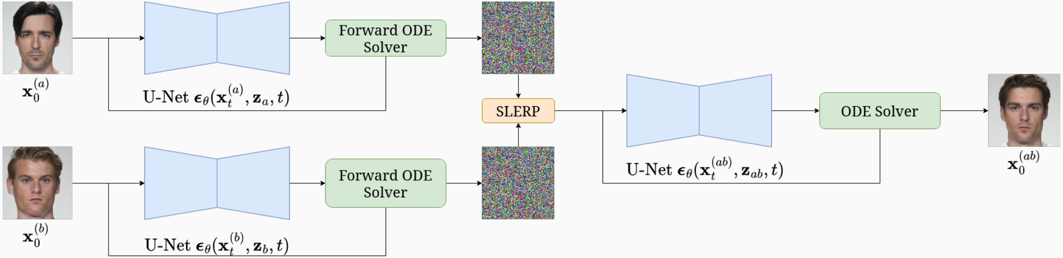


- Let $\Phi(\mathbf{x}_0, \mathbf{z}, \boldsymbol{f}_\theta, \{t_n\}_{n=1}^N) \mapsto \mathbf{x}_T$ denote a numerical ODE solver with:
  1. Initial image $\mathbf{x}_0$,
  2. Latent representation of $\mathbf{x}_0$, $\mathbf{z} = E(\mathbf{x}_0)$,
  3. Denoising U-Net conditioned on $\mathbf{z}$, $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t,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

$$\boldsymbol{f}_\theta(\mathbf{x}_t, \mathbf{z}, t) = f(t)\mathbf{x}_t + \frac{g^2(t)}{2\sigma_t}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{z}, t), \tag{6}$$

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

Face Morphing with Diffusion

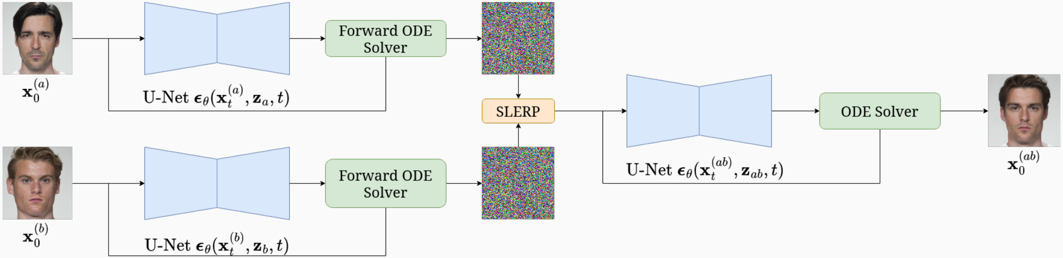


- Encode images solving the PF ODE as time runs *forwards*:

$$\mathbf{x}_T^{(\{a,b\})} = \Phi(\mathbf{x}_0^{(\{a,b\})}, \mathbf{z}_{\{a,b\}}, \boldsymbol{f}_\theta, \{t_n\}_{n=1}^{N_F}), \tag{7}$$

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

Face Morphing with Diffusion

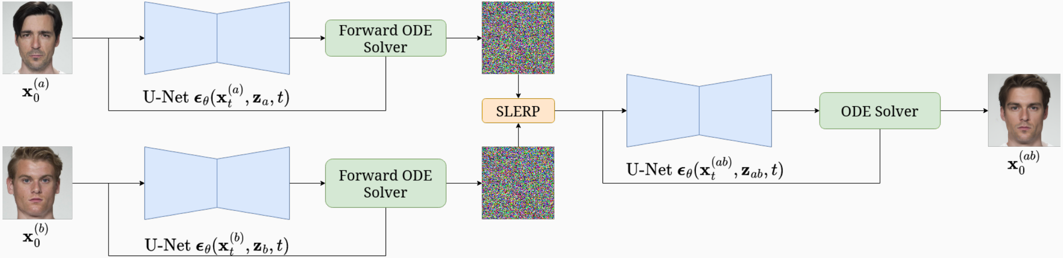


- Morph the latent representations:

$$\mathbf{x}_T^{(ab)} = \operatorname{slerp}(\mathbf{x}_T^{(a)}, \mathbf{x}_T^{(b)}; \gamma), \tag{8}$$

$$\mathbf{z}_{ab} = \operatorname{lerp}(\mathbf{z}_a, \mathbf{z}_b; \gamma), \tag{9}$$

by a factor of $\gamma = 0.5$.

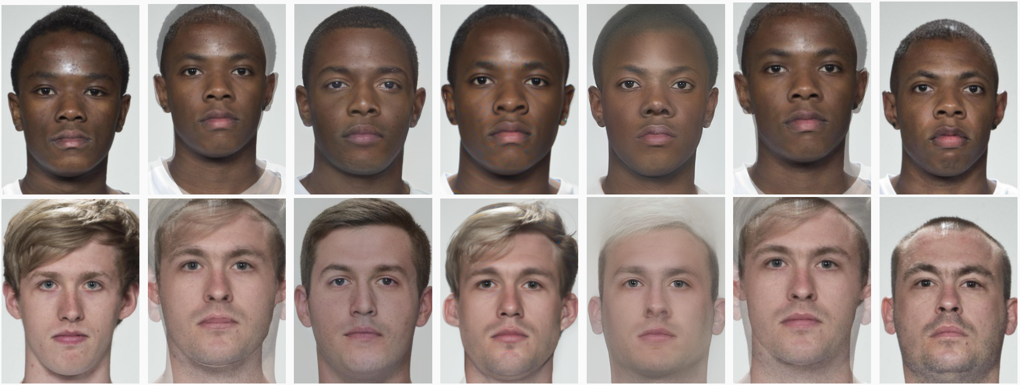


- Create morph by solving the PF ODE as time runs *backwards*:

$$\mathbf{x}_0^{(ab)} = \Phi(\mathbf{x}_T^{(ab)}, \mathbf{z}_{ab}, \boldsymbol{f}_\theta, \{\tilde{t}_n\}_{n=1}^N), \tag{10}$$

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 (↑)		
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) [11]:

$$M(\delta) = \frac{1}{M}\sum_{n=1}^{M}\left\{\left[\min_{n\in\{1,\ldots,N_m\}} S_m^n\right] > \delta\right\}, \tag{11}$$

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

Summary

- DiM creates morphs with high visual fidelity.
- DiM outperforms GAN-based morphs.
- DiM is difficult to detect if not explicitly trained against.
- Our article "Leveraging Diffusion For Strong and High Quality Face Morphing Attacks" was accepted in IEEE TBIOM[3].

[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.

Greedy-DiM

Guided Optimization for Morphing

- MIPGAN[4] showed the power in using guided optimization for face morphing.
- MIPGAN far outperforms the unguided GAN architecture.
- Can we do this for DiMs?
- It is difficult to find the optimal $\mathbf{x}_T^{(ab)}$ and $\mathbf{z}_{ab}$ in DiMs.
- Morph-PIPE solves this via brute force search[5].
- Can we do better?

[4] 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.

[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.

Yes, by being greedy

Identity Guided Morphing

Table 3: Comparison of existing DiM methods in the literature and our proposed algorithm.

	DiM [3]	**Fast-DiM [1]**	**Morph-PIPE [14]**	**Ours (Greedy-DiM)**
ODE solver	DDIM	DPM++ 2M	DDIM	DDIM
Forward ODE solver	DiffAE	DDIM	DiffAE	DiffAE
Number of sampling steps	100	50	2100	20
Heuristic function	✗	✗	$\mathcal{L}^*_{ID}$	$\mathcal{L}^*_{ID}$
Search strategy	✗	✗	Brute-force search	Greedy optimization
Search space	$\emptyset$	$\emptyset$	Set of 21 blend values	Image space
Optimal solution in search space	✗	✗	0	1

- Greedily search for the optimal ϵ at each time step which minimizes the identity loss defined as the sum of two sub-losses:

$$\mathcal{L}_{ID} = d(v_{ab}, v_a) + d(v_{ab}, v_b), \tag{12}$$

$$\mathcal{L}_{diff} = \Big|d(v_{ab}, v_a) - d(v_{ab}, v_b))\Big|, \tag{13}$$

$$\mathcal{L}^*_{ID} = \mathcal{L}_{ID} + \mathcal{L}_{diff}, \tag{14}$$

where $v_a = F(\mathbf{x}_0^{(a)}), v_b = F(\mathbf{x}_0^{(b)}), v_{ab} = F(\mathbf{x}_0^{(ab)})$, and $F : \mathcal{X} \to V$ is an FR system which embeds images into a vector space V which is equipped with a measure of distance, d.

Greedy-DiM*

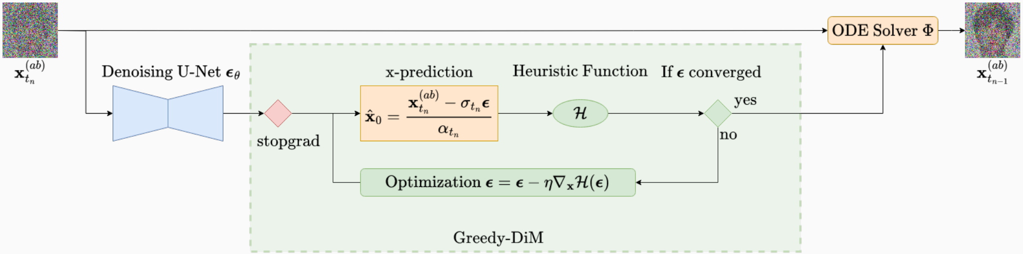


Figure 4: Overview of a single step of the Greedy-DiM* algorithm. Proposed changes highlighted in green.

- During each step greedily solve for the best predicted noise, ϵ.

Greedy-DiM*

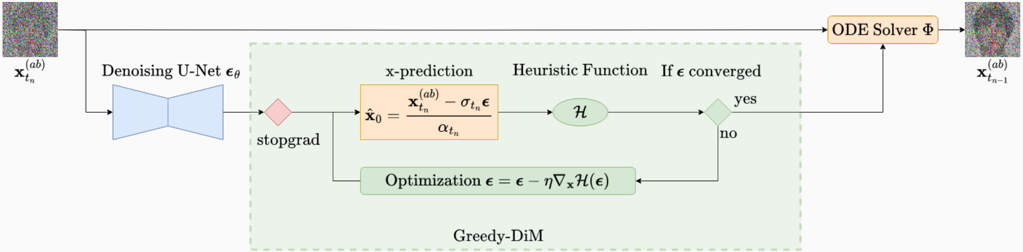


Figure 4: Overview of a single step of the Greedy-DiM* algorithm. Proposed changes highlighted in green.

- Take prediction from model $\boldsymbol{\epsilon} = \mathrm{stopgrad}(\boldsymbol{\epsilon}_\theta(\mathbf{x}_t^{(ab)}, \mathbf{z}_{ab}, t))$.

Greedy-DiM*

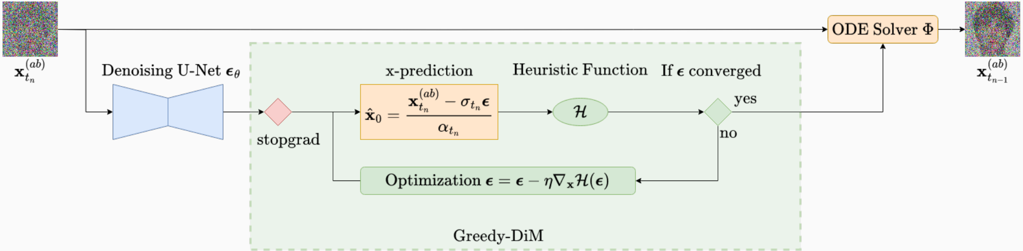


Figure 4: Overview of a single step of the Greedy-DiM* algorithm. Proposed changes highlighted in green.

- Perform a one-shot prediction of $\mathbf{x}_0$ via:

$$\hat{\mathbf{x}}_0 = \frac{\mathbf{x}_t^{(ab)} - \sigma_t \boldsymbol{\epsilon}}{\alpha_t}. \tag{15}$$

Greedy-DiM*

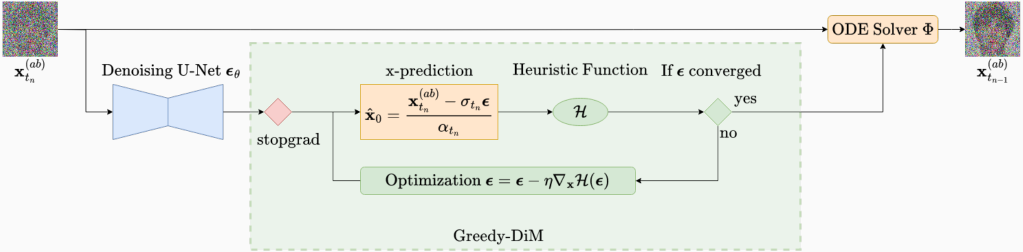


Figure 4: Overview of a single step of the Greedy-DiM* algorithm. Proposed changes highlighted in green.

- Perform gradient descent on $\boldsymbol{\epsilon}$ via:

$$\boldsymbol{\epsilon} = \boldsymbol{\epsilon} - \eta\nabla_{\mathbf{x}}\mathcal{H}(\hat{\mathbf{x}}_0). \tag{16}$$

- Use the optimal $\boldsymbol{\epsilon}^*$ to then find the next step $\mathbf{x}_s^{(ab)}$, $s < t$.

Visual Results

(a) Identity a (b) DiM-A (c) Fast-DiM (d) Morph-PIPE (e) Greedy-DiM-S (f) Greedy-DiM* (g) Identity b

Figure 5: Comparison of DiM morphs on the FRLL dataset.

Results

Table 4: Vulnerability of different FR systems across different morphing attacks on the SYN-MAD 2022 dataset. FMR = 0.1%.

		MMPMR(↑)		
Morphing Attack	NFE (↓)	**AdaFace** [8]	**ArcFace** [6]	**ElasticFace** [4]
FaceMorpher [7]	-	89.78	87.73	89.57
Webmorph [7]	-	97.96	96.93	98.36
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]	350	92.23	90.18	93.05
Fast-DiM [1]	300	92.02	90.18	93.05
Fast-DiM-ode [1]	150	91.82	88.75	91.21
Morph-PIPE [14]	2350	95.91	92.84	95.5
Greedy-DiM* [2]	270	**100**	**100**	**100**

Results

Table 5: Vulnerability of different FR systems across different morphing attacks on the SYN-MAD 2022 dataset. FMR = 0.01%.

		MMPMR(↑)		
Morphing Attack	NFE (↓)	**AdaFace** [8]	**ArcFace** [6]	**ElasticFace** [4]
FaceMorpher [7]	-	66.05	64.01	70.96
Webmorph [7]	-	77.3	79.55	85.69
OpenCV [7]	-	58.9	62.58	71.98
MIPGAN-I [13]	-	15.75	23.52	21.88
MIPGAN-II [13]	-	11.04	19.22	17.79
DiM [3]	350	58.9	58.69	67.28
Fast-DiM [1]	300	55.83	55.42	65.85
Fast-DiM-ode [1]	150	54.19	53.58	63.8
Morph-PIPE [14]	2350	62.37	61.76	71.78
Greedy-DiM* [2]	270	**85.89**	**91.62**	**96.11**

Summary

- SOTA performance on SYN-MAD 2022 dataset.
- Adds only a little overhead to vanilla DiM.
- Guiding heuristic $\mathcal{H}$ can be swapped for another differentiable function.
- Our paper "Greedy-DiM: Greedy Algorithms for Unreasonably Effective Face Morphs" was accepted at IJCB 2024[6].

[6] Zander W. Blasingame and Chen Liu. "Greedy-DiM: Greedy Algorithms for Unreasonably Effective Face Morphs". In: *2024 IEEE International Joint Conference on Biometrics (IJCB)*. Sept. 2024, pp. 1–10.

Questions?

Code and project page for Greedy-DiM

Further reading about DiM models

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: *2024 IEEE International Joint Conference on Biometrics (IJCB)*. Sept. 2024, pp. 1–10.

[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.

[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] Ulrich Scherhag et al. "Biometric Systems under Morphing Attacks: Assessment of Morphing Techniques and Vulnerability Reporting". In: *2017 International Conference of the Biometrics Special Interest Group (BIOSIG)*. 2017, pp. 1–7. DOI: 10.23919/BIOSIG.2017.8053499.

[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.