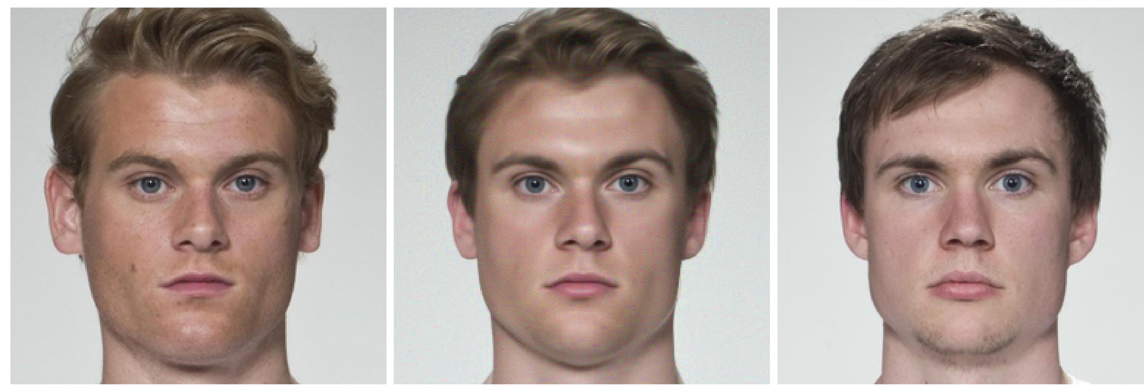
Greedy-DiM: Greedy Algorithms for Unreasonably Effective Face Morphs

Zander W. Blasingame Chen Liu
Department of Electrical and Computer Engineering
Clarkson University
{blasinzw, cliu}@clarkson.edu

IEEE International Joint Conference on Biometrics

Motivation



(a) identity a (b) Morphed image (c) identity b

Figure 1. Example of a morph created using Greedy-DiM. Samples are from the FRLL dataset [1].

- **Di**ffusion **M**orphs **(DiM)** are a recent SOTA algorithm for creating face morphs [2]
- Identity guided generation greatly increases the effectiveness of face morphing [3]
- Currently, there exists *no* algorithm for DiMs which perform identity *guided* generation!
- We propose Greedy-DiM, a family of algorithms to perform identity guided generation with diffusion models

Table 1. Comparison of existing DiM methods in the literature and our proposed algorithm.

	DiM [2]	Fast-DiM [4]	Morph-PIPE [5]	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 ($\mathcal{S}$)	✗	✗	21 Morphs	Image space ($\mathcal{X}$)
$\mathbb{P}(\mathcal{S})$	✗	✗	0	1

Methodology

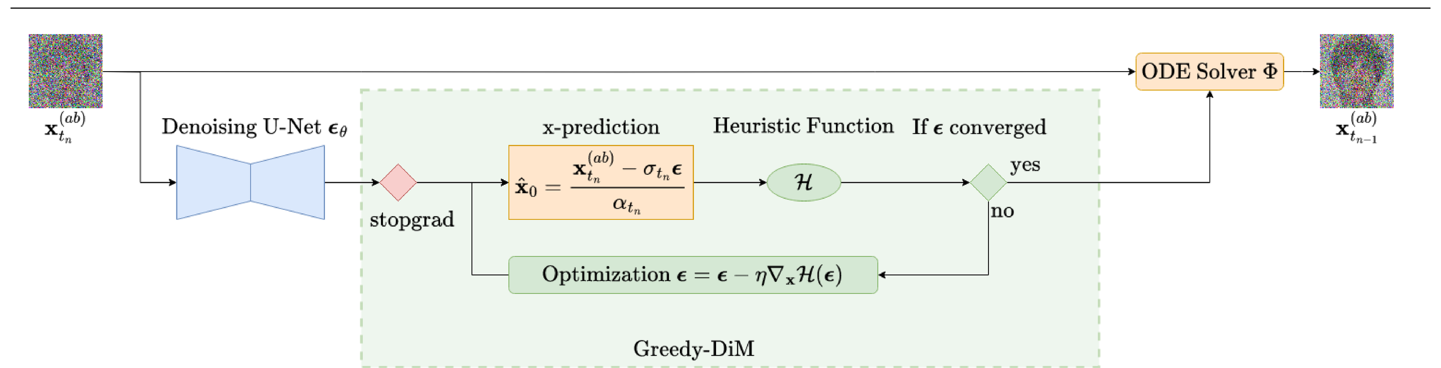


Figure 2. Overview of a single step of the Greedy-DiM* algorithm. Proposed changes highlighted in green.

- The Variance Preserving (VP) diffusion process is governed by an Itô SDE

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

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

with noise schedule $\alpha_t^2 + \sigma_t^2 = 1$ such that $\mathbf{x}_t = \alpha_t\mathbf{x}_0 + \sigma_t\boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ [6]
- Diffusion models train a U-Net to learn the added noise $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) \approx \boldsymbol{\epsilon}$
- To draw samples from $p_{data}(\mathbf{x}) = p_0(\mathbf{x})$, solve the *Probability Flow* ODE [6]

$$\frac{\mathrm{d}\mathbf{x}_t}{\mathrm{d}t} = f(t)\mathbf{x}_t + \frac{g^2(t)}{2\sigma_t}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) \tag{3}$$

- Let Φ denote a first-order numerical ODE solver to the PF ODE
- We use the identity loss $\mathcal{L}_{ID}^*$ [3] defined as

$$\mathcal{L}_{ID} = d(v_{ab}, v_a) + d(v_{ab}, v_b) \qquad \mathcal{L}_{diff} = \left|d(v_{ab}, v_a) - d(v_{ab}, v_b)\right| \tag{4}$$

$$\mathcal{L}_{ID}^* = \mathcal{L}_{ID} + \mathcal{L}_{diff} \tag{5}$$

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
- Greedily search for optimal $\boldsymbol{\epsilon}^*$ w.r.t $\mathcal{H}$ at each time step t_n using $\mathbf{x}_0$-prediction

$$\hat{\mathbf{x}}_0 = \frac{\mathbf{x}_{t_n}^{(ab)} - \sigma_t\boldsymbol{\epsilon}}{\alpha_t} \tag{6}$$

- **Greedy-DiM-S**: Preforms a greedy search over 21 blend values of $\boldsymbol{\epsilon}$ at each step t_n
- **Greedy-DiM***: Greedy gradient descent over $\mathcal{X}$ to find $\boldsymbol{\epsilon}^*$

Highlighted Results

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

Figure 3. Comparison of DiM morphs on the FRLL dataset.

- Evaluated the proposed morphing attack on the recent SYN-MAD 2022 dataset [7]
- Compared against three landmark-based morphs: OpenCV, FaceMorpher, and Webmorph
- Compared against two identity GAN algorithms: MIPGAN-I and MIPGAN-II
- Compared against prior DiM algorithms: DiM-A, DiM-C, Fast-DiM, Fast-DiM-ode, and Morph-PIPE
- Used three FR systems representing the SOTA: ArcFace [8], AdaFace [9], and ElasticFace [10]
- The Mated Morph Presentation Match Rate (MMPMR) metric [11] is defined as

$$M(\delta) = \frac{1}{M}\sum_{n=1}^{M}\left\{\left[\min_{n\in\{1,\ldots,N_m\}} S_m^n\right] > \delta\right\} \tag{7}$$

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
- The Morphing Attack Potential (MAP) [12] metric is defined such that $\mathrm{MAP}[r, c]$ denotes the proportion of morphed images that successfully register a false accept against at least r attempts against each contributing subject of at least c FR systems

Table 2. Vulnerability of different FR systems across different morphing attacks on the SYN-MAD 2022 dataset. FMR = 0.1%.

		MMPMR(↑)		
Morphing Attack	NFE(↓)	AdaFace	ArcFace	ElasticFace
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 [3]	-	72.19	77.51	66.46
MIPGAN-II [3]	-	70.55	72.19	65.24
DiM-A [2]	350	92.23	90.18	93.05
DiM-C [2]	350	89.57	83.23	86.3
Fast-DiM [4]	300	92.02	90.18	93.05
Fast-DiM-ode [4]	150	91.82	88.75	91.21
Morph-PIPE [5]	2350	95.91	92.84	95.5
Greedy-DiM-S	350	95.71	93.87	95.3
Greedy-DiM*	270	**100**	**100**	**100**

Table 3. MAP(↑) metric for all three FR systems on the SYN-MAD 2022 dataset. FMR = 0.1%.

		Number of FR Systems		
Morphing Attack	NFE(↓)	1	2	3
FaceMorpher [7]	-	92.23	89.57	85.28
Webmorph [7]	-	98.77	98.36	96.11
OpenCV [7]	-	97.55	93.87	89.78
MIPGAN-I [3]	-	85.07	72.39	58.69
MIPGAN-II [3]	-	80.37	69.73	57.87
DiM-A [2]	350	96.93	92.43	86.09
DiM-C [2]	350	92.84	87.53	78.73
Fast-DiM [4]	300	97.14	92.43	85.69
Fast-DiM-ode [4]	150	95.91	91.21	84.66
Morph-PIPE [5]	2350	98.16	95.71	90.39
Greedy-DiM-S	350	97.34	95.71	91.82
Greedy-DiM*	270	**100**	**100**	**100**

Theoretical Results

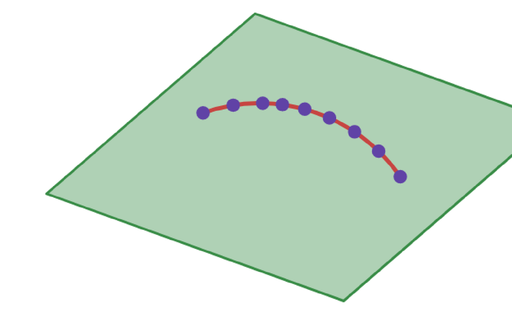


Figure 4. Illustration of the search space in $\mathbb{R}^2$ of different DiM algorithms at a single step. Purple denotes Morph-PIPE/Greedy-DiM-S, red denotes Greedy-DiM-S continuous, and green denotes Greedy-DiM*.

Theorem 1. *Given a sequence of monotonically descending time steps, $\{t_n\}_{n=1}^N$, from T to 0, the DDIM solver to the Probability Flow ODE, and a heuristic function $\mathcal{H}$, then the locally optimal solution admitted by Greedy-DiM* at time t_n is globally optimal.*

Theorem 2. *Let $\mathbb{P}$ be a probability distribution on a compact subset $\mathcal{X} \subseteq \mathbb{R}^n$ with full support on $\mathcal{X}$ which models the distribution of the optimal $\mathbf{x}_0^*$ and is absolutely continuous w.r.t. the n-dimensional Lebesgue measure λ^n on $\mathcal{X}$. Let $\mathcal{S}_P, \mathcal{S}_S, \mathcal{S}^*$ denote the search spaces of the Morph-PIPE, Greedy-DiM-S, and Greedy-DiM* algorithms. Then the following statements are true.*

1. $\mathbb{P}(\mathcal{S}_P) = \mathbb{P}(\mathcal{S}_S) = 0.$
2. $\mathbb{P}(\mathcal{S}^*) = 1.$

Additional Morphed Images

Figure 5. Morphed images generated via Greedy-DiM*.

Conclusion

- SOTA morphing attack which outperforms all previous morphing attacks
- First representation-based morphing attack to *consistently* outperform landmark-based morphing attacks
- Developed a novel strategy to incorporate identity guidance for diffusion models
- Adds little overhead compared to the original DiM algorithms
- Much less overhead than Morph-PIPE with superior performance
- Greedy guided generation can be applied to other guided diffusion problems

References

[1] L. DeBruine and B. Jones, "Face Research Lab London Set," 5 2017.
[2] Z. W. Blasingame and C. Liu, "Leveraging diffusion for strong and high quality face morphing attacks," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 6, no. 1, pp. 118–131, 2024.
[3] H. Zhang, S. Venkatesh, R. Ramachandra, K. Raja, N. Damer, and C. Busch, "Mipgan—generating strong and high quality morphing attacks using identity prior driven gan," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 3, no. 3, pp. 365–383, 2021.
[4] Z. W. Blasingame and C. Liu, "Fast-dim: Towards fast diffusion morphs," *IEEE Security & Privacy*, pp. 2–13, 2024.
[5] H. Zhang, R. Ramachandra, K. Raja, and B. Christoph, "Morph-pipe: Plugging in identity prior to enhance face morphing attack based on diffusion model," in *Norwegian Information Security Conference (NISK)*, 2023.
[6] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," in *International Conference on Learning Representations*, 2021.
[7] M. Huber, F. Boutros, A. T. Luu, K. Raja, R. Ramachandra, N. Damer, P. C. Neto, T. Gonçalves, A. F. Sequeira, J. S. Cardoso, J. Tremoço, M. Lourenço, S. Serra, E. Cermeño, M. Ivanovska, B. Batagelj, A. Kronovšek, P. Peer, and V. Štruc, "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)*, pp. 1–10, 2022.
[8] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4690–4699, 2019.
[9] M. Kim, A. K. Jain, and X. Liu, "Adaface: Quality adaptive margin for face recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.
[10] F. Boutros, N. Damer, F. Kirchbuchner, and A. Kuijper, "Elasticface: Elastic margin loss for deep face recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 1578–1587, June 2022.
[11] U. Scherhag, A. Nautsch, C. Rathgeb, M. Gomez-Barrero, R. N. J. Veldhuis, L. Spreeuwers, M. Schils, D. Maltoni, P. Grother, S. Marcel, R. Breithaupt, R. Ramachandra, and C. Busch, "Biometric systems under morphing attacks: Assessment of morphing techniques and vulnerability reporting," in *2017 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pp. 1–7, 2017.
[12] M. Ferrara, A. Franco, D. Maltoni, and C. Busch, "Morphing attack potential," in *2022 International Workshop on Biometrics and Forensics (IWBF)*, pp. 1–6, 2022.

Supported by CITeR and NSF under Grant # 1650503