Poster: Diffusion Based Face Generation via Image Editing and Image Morphing

Liyue Fan, Joseph Roberson

UNC Charlotte

Liyue.fan@charlotte.edu, jrobe187@charlotte.edu

Abstract

AI-generated content has important implications on creative work and potential consequences in society. The detection of AI-generated content has been intensively studied by the research community. Prior research shows that curating datasets of AI-generated content requires significant efforts and can advance the development of detection methods. Recently released AI content generation datasets focus on face images and are generated by state-of-the-art synthesis models, such as diffusion models. In this study, we propose to extend existing AI-generated face datasets by incorporating recent results in image editing and image morphing in the latent space of diffusion models. Furthermore, we quantify the quality of the newly generated data using widely adopted quality measures for generative models and for facial image quality assessment; we also evaluate the performance of a state-of-the-art fake image detection method on the generated datasets. Our quantitative and qualitative results show that data generated by image editing and image morphing provide interesting additions to existing face forgery datasets. The generated data can be found at: https://github.com/fan-group/LatentFake.

Acknowledgement

This work has been accepted by the IEEE CIC/CogMI/TPS 2025 Joint Conference in November 2025. The original abstract and author list are shown above. Conference program can be found at:

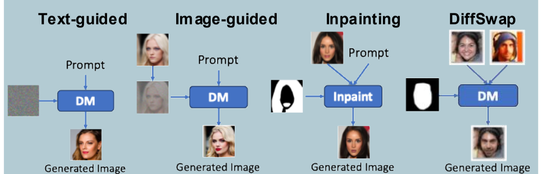
https://www.sis.pitt.edu/lersais/conference/tps/2025/technical-program/

Diffusion Based Face Generation via Image Editing and Image Morphing

Liyue Fan PhD, Joseph Roberson
Liyue.fan@charlotte.edu, jrob187@charlotte.edu
University of North Carolina at Charlotte

Introduction

- Malicious applications of DeepFake techniques may lead to grave consequences.
- Significant efforts have been made to organize DeepFake detection challenges and to curate datasets [1-4]
- **Conditional generation** informs the data synthesis process with input conditions, e.g., text prompts.
- Recent benchmark datasets [4] with the following diffusion based conditional generations:



Opportunities:

1. Recent results in latent space editing and interpolation [5,6] may produce new DeepFake datasets
2. Evaluation on DeepFake detection and generative quality may quantify the usefulness of DeepFake datasets

Observations

- Image editing and image morphing generate visually plausible synthetic images and reasonable quality metrics.
- These newly generated datasets may reveal new challenges to DeepFake detection methods.
- Future work may conduct comprehensive analysis to understand the classification of real vs. generated distributions.

Acknowledgement: LF and JR are supported in part by the National Science Foundation under grant CNS-2027114. The opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

Generation Method

Image Editing

- Recent work [5] proposes x-space guidance:

$$\tilde{x}_t = x_t + \gamma[\epsilon_\theta(x_t + v) - \epsilon_\theta(x_t)]$$

- Images can be edited in the latent space of diffusion models along any local basis vectors, e.g., v.

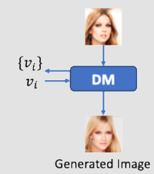
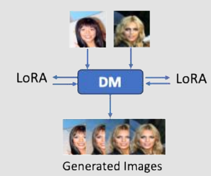


Image Morphing

- [6] generates smooth transition between two input images.
- LoRA parameters and latent noises are interpolated.



Usefulness Evaluation

Image Editing samples:

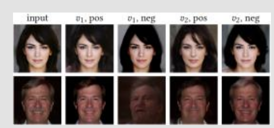
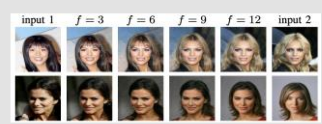


Image Morphing samples:



DeepFake Detection

- Classifier [7] on real and fake images with CLIP-ViT model

Generative Quality

- FID: distributional similarity between real and fake images
- Improved Precision and Recall (IPR): trade-off between sample quality and manifold coverage

Results

DeepFake Detection on generated datasets

- Model is trained with real and stable diffusion image-guided generation images, and tested on real and IM or IE images

Dataset	TNR ⇑	TPR ⇑	Accuracy ⇑	AUC ⇑
IM $s = 3$	0.001	0.999	0.500	0.366
IM $s = 6$	0.005	0.998	0.501	0.437
IM $s = 9$	0.007	0.995	0.501	0.437
IM $s = 12$	1.000	0.000	0.500	0.368
IE v_1,pos	0.851	0.815	0.833	0.914
IE v_1,neg	0.851	0.820	0.835	0.913
IE v_2,pos	0.872	0.846	0.859	0.935
IE v_2,neg	0.874	0.826	0.850	0.925

Quality of generated datasets

- Metrics are computed between real input images and IM or IE images

Dataset	FID ⇓	Precision ⇑	Recall ⇑
IM $s = 3$	18.569	0.732	0.381
IM $s = 6$	41.100	0.637	0.180
IM $s = 9$	41.184	0.626	0.183
IM $s = 12$	18.860	0.739	0.373
IE v_1,pos	35.145	0.774	0.418
IE v_1,neg	28.253	0.75	0.417
IE v_2,pos	30.865	0.762	0.420
IE v_2,neg	29.337	0.774	0.435

⇑: higher is better; ⇓: lower is better

References

1. A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "Faceforensics++: Learning to detect manipulated facial images," in Proceedings of the IEEE/CVF international conference on computer vision, 2019, pp. 1–11.
2. Y. Li, X. Yang, P. Sun, H. Qi, and S. Lyu, "Celeb-df: A large-scale challenging dataset for deepfake forensics," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 3207–3216.
3. B. Dolhansky, J. Bitton, B. Pflaum, J. Lu, R. Howes, M. Wang, and C. C. Ferrer, "The deepfake detection challenge (dfdc) dataset," arXiv preprint arXiv:2006.07397, 2020.
4. Z. Chen, K. Sun, Z. Zhou, X. Lin, X. Sun, L. Cao, and R. Ji, "Diffusionface: Towards a comprehensive dataset for diffusion-based face forgery analysis," arXiv preprint arXiv:2403.18471, 2024.
5. Y.-H. Park, M. Kwon, J. Choi, J. Jo, and Y. Uh, "Understanding the latent space of diffusion models through the lens of Riemannian geometry," Advances in Neural Information Processing Systems, vol. 36, pp. 24 129–24 142, 2023.
6. K. Zhang, Y. Zhou, X. Xu, B. Dai, and X. Pan, "Diffmorpher: Un-leashing the capability of diffusion models for image morphing," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 7912–7921.
7. U. Ojha, Y. Li, and Y. J. Lee, "Towards universal fake image detectors that generalize across generative models," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 24 480–24 489.
8. T. Kynk¨a¨anniemi, T. Karras, S. Laine, J. Lehtinen, and T. Aila, "Improved precision and recall metric for assessing generative models," Advances in Neural Information Processing Systems, vol. 32, 2019.