Technische Universität München – Faculty of Informatics
Chair for Computer Aided Medical Procedures (Prof. Nassir Navab)
Practical Course: Machine Learning in Medical Imaging WS24/25

Conditional Facial Image Generation and Evaluation via Compositional Profiling

1 General Info

Project Title: Conditional Facial Image Generation and Evaluation via Compositional Profiling
Supervisors: Azade Farshad, Yousef Yeganeh
Contact Email: azade.farshad@tum.de, y.yeganeh@tum.de

2 Background and Motivation

Accurate facial image generation conditioned on disease profiles can significantly impact healthcare, particularly in generating new data based on medical textbooks and visualizing disease progression while maintaining privacy. Generating realistic images that reflect facial changes due to disease progression remains challenging due to the limited availability of training data and privacy concerns [6]. Incorporating structured medical information—such as textbooks, tabular data, and knowledge graphs (KGs) [2]—alongside advanced generative techniques could create new possibilities for early diagnosis and disease monitoring [8].

3 Project Outline

Profiles Collection and Preparation The following sources or similar ones can be utilized for the extraction of the profiles: (1) Healthy face datasets such as CelebA [5], (2) Diseased face datasets such as DSF [3], (3) Literature and medical knowledge graphs [2] and large language models [9] for structured information extraction; these profiles can then be transformed into prompts, tabular data, masks, or meshes.

Generative Model Conditioning Using backbones of generative models like Stable Diffusion [7] and state-of-the-art conditioning mechanisms such as IP-Adapter [10], LEDITS++ [1], ControlNet or similar methods [4], we will condition the models on disease profiles to generate realistic facial images reflecting different disease stages.

Evaluation The evaluation could focus on: (a) **Identity Consistency**: Ensuring the generated images preserve the individual's identity using facial recognition metrics, (b) **Disease-Specific Characteristics**: Assessing the accuracy of disease-specific features by comparing with clinical descriptions and real-world images.

4 Project Objectives

The primary goal of this project is to enhance the evaluation and accuracy of generative models for disease-specific facial image synthesis. The specific objectives are:

- To develop advanced conditioning mechanisms for generative models using compositional facial profiling.
- To evaluate the generated images for identity consistency and disease-specific facial characteristics.
- To generate consistent video frames capturing the identity and disease profile.

5 Technical Prerequisites

- Strong foundation in machine learning and deep learning.
- Experience with PyTorch and Python.
- Proficiency with generative models.

6 Benefits

- Leverage our existing works on the topic and prior experience to accelerate your progress.
- Regular supervision and feedback.
- Potential for novel research contributions.
- Opportunity to publish findings in conferences or journals.

WP	Description
WP1	**Literature Review**: Familiarization with disease-specific facial features, generative models, and conditioning mechanisms.
WP2	**Data Preparation**: Collecting and preprocessing datasets; extracting features from medical texts and KGs.
WP3	**Baseline Implementation**: Implementing generative baselines for disease-specific facial synthesis.
WP4	**Advanced Conditioning**: Developing and implementing advanced conditioning methods for generative models.
WP5	**Evaluation Development**: Designing evaluation metrics for identity consistency and disease-specific accuracy.
WP6	**Video Generation**: Generating consistent video frames capturing identity and disease progression.
WP7	**Finalization and Reporting**: Finalizing results, conducting evaluations, and preparing documentation.

Table 1: Suggested Work Packages

References

[1] Manuel Brack, Felix Friedrich, Katharia Kornmeier, Linoy Tsaban, Patrick Schramowski, Kristian Kersting, and Apolinário Passos. Ledits++: Limitless image editing using text-to-image models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8861–8870, 2024.

[2] Payal Chandak, Kexin Huang, and Marinka Zitnik. Building a knowledge graph to enable precision medicine. *Nature Scientific Data*, 2023.

[3] Bo Jin. Disease-specific faces, 2020.

[4] Feng-Li Lian, James R Moyne, and Dawn M Tilbury. Performance evaluation of control networks: Ethernet, controlnet, and devicenet. *IEEE control systems magazine*, 21(1):66–83, 2001.

[5] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Large-scale celebfaces attributes (celeba) dataset. *Retrieved August*, 15(2018):11, 2018.

[6] Anwesha Mohanty, Alistair Sutherland, Marija Bezbradica, and Hossein Javidnia. High fidelity synthetic face generation for rosacea skin condition from limited data. *arXiv preprint arXiv:2303.04839*, 2023.

[7] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695, 2022.

[8] Kuan Wang and Jiebo Luo. Detecting visually observable disease symptoms from faces. *EURASIP Journal on Bioinformatics and Systems Biology*, 2016:1–8, 2016.

[9] Michihiro Yasunaga, Antoine Bosselut, Hongyu Ren, Xikun Zhang, Christopher D Manning, Percy S Liang, and Jure Leskovec. Deep bidirectional language-knowledge graph pretraining. *Advances in Neural Information Processing Systems*, 35:37309–37323, 2022.

[10] Hu Ye, Jun Zhang, Sibo Liu, Xiao Han, and Wei Yang. Ip-adapter: Text compatible image prompt adapter for text-to-image diffusion models. *arXiv preprint arXiv:2308.06721*, 2023.