FLAME GAN: High-Fidelity 3D Face Reconstruction via Disentangled Representations and Explicit Geometric Control

Wiem Grina and Ali Douik

Electronic Engineering, National Engineering School in Monastir, University of Monasti, Tunisia

Networked Objects Control and Communication Systems Laboratory, Sousse Technology Center, National School of Engineers of Sousse, University of Sousse

ABSTRACT

FLAME GAN is a new framework that reconstructs high-fidelity 3D faces from a single 2D image. By combining **GANs** for realism with the **FLAME parametric model** for control, our system disentangles identity, expression, and pose. The use of **3DMM priors** and a tailored loss function ensures accuracy, resulting in superior performance in generating realistic and consistent 3D faces.

PROBLEM

The challenge of reconstructing a high-fidelity 3D face from just one 2D image.

SOLUTION

- To introduce a new framework called FLAME GAN for high-fidelity and controllable 3D face reconstruction from a single 2D image.
- To support identity-aware retrieval and similarity-based applications.
- To combine the strengths of Generative Adversarial Networks (GANs), the FLAME parametric 3D face model, and 3D Morphable Model (3DMM) priors to achieve a robust and interpretable system.

METHODS

- The core of our approach lies in extending the DiscoFaceGAN architecture to disentangle identity, expression, and pose.
- The model utilizes a comprehensive loss function with five key components: an adversarial loss (L_{adv}) for photorealism, an identity loss (L_{id}) to preserve facial identity, a reconstruction loss (L_{rec}) for image fidelity, a geometric consistency loss (L_{geo}) to ensure accurate 3D shape, and a semantic parsing loss (L_{CE}) to guide the network with a facial mask.
- This robust formulation allows us to reconstruct high-fidelity 3D faces from a single 2D image while maintaining precise control over key facial parameters.

$$L_{total} = \lambda_{adv}L_{adv} + \lambda_{id}L_{id} + \lambda_{rec}L_{rec} + \lambda_{geo}L_{geo} + \lambda_{CE}L_{CE}$$

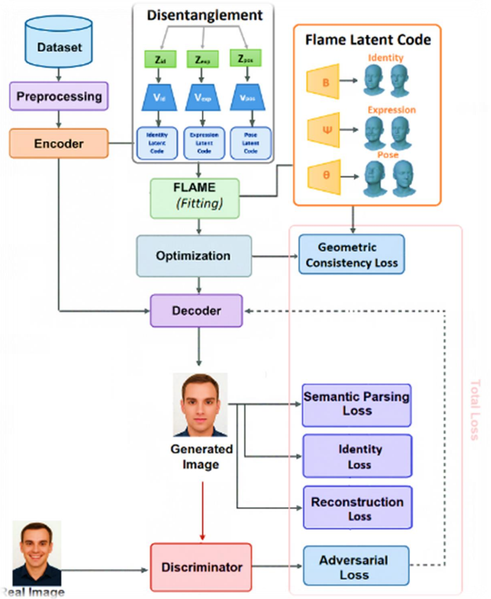

Fig1. FLAME GAN Framework Overview.

RESULTS

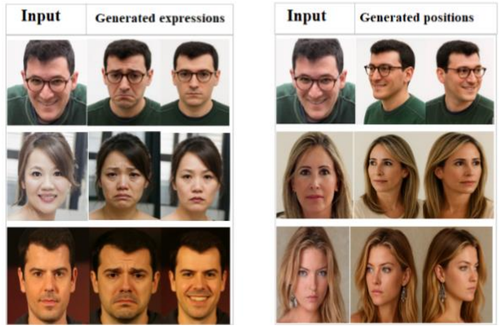

Fig 2. Examples of Synthesized Facial Expressions and Head Poses.

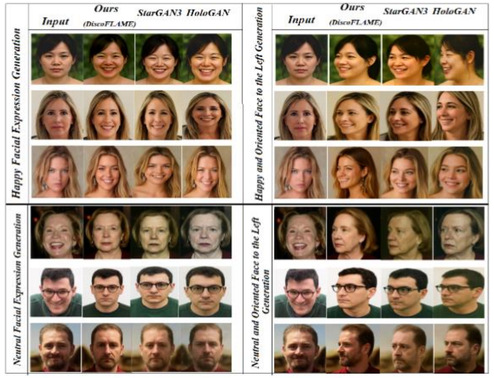

Fig3. Qualitative Comparison of FLAME GAN with StarGAN3 and HoloGAN for Facial Expression and Orientation Generation.

Experiments

Table 1 Quantitative Comparison: FID Score (lower is better), Identity Similarity (Cosine, higher is better), and LPIPS Score (lower is better

Method	FID Score	Identity Similarity	LPIPS Score
DiscoFaceGAN	10.12	0.880	0.285
StyleGAN3	8.42	0.812	0.220
StarGAN	13.55	0.768	0.350
HoloGAN	14.87	0.740	0.380
StyleFlow	9.75	0.861	0.270

Table 2 Identity Preservation Evaluation for Our FLAME GAN Model Across Expression and Pose Changes (Higher is Better)

Attribute	VGG-Face	ArcFace	FaceNet
Expression	0.91	0.95	0.93
Pose	0.90	0.92	0.89

Table 3 Chamfer Distance (↓) between predicted and ground-truth 3D meshes (×10−3)

Method	Chamfer Distance
DiscoFaceGAN	2.82
StyleGAN3	3.42
StarGAN	4.01
HoloGAN	4.58
StyleFlow	2.95
Ours (DiscoFaceGAN + FLAME)	**2.35**

Table 4 Ablation Study Results for FLAME GAN

Experiment	Component Removed	SSIM	MSE	Diversity Score
Baseline	None	0.95	0.02	0.90
w/o L_{adv}	Generative Adversarial Loss	0.86	0.07	0.81
w/o L_{id}	Identity Loss	0.90	0.03	0.87
w/oL_{rec}	Reconstruction Loss	0.91	0.04	0.85
w/o L_{geo}	Geometric Consistency Loss	0.89	0.05	0.84
w/o $L_{parsing}$	Semantic Parsing Loss	0.92	0.03	0.88
No Expression Control	Expression Control	0.87	0.06	0.82
No Posture Control	Posture Control	0.88	0.06	0.83

CONCLUSION

FLAME GAN is a framework that reconstructs high-fidelity 3D faces from a single image by combining GANs and the FLAME parametric model. It provides disentangled control over identity, expression, and pose, and achieves superior performance guided by 3DMM priors.