

StyleGene: Crossover and Mutation of Region-level Facial Genes for Kinship Face Synthesis

Hao Li¹, Xianxu Hou^{2,5}, Zepeng Huang¹, Linlin Shen^{1,2,3,4*}

¹Computer Vision Institute, College of Computer Science and Software Engineering, Shenzhen University

²National Engineering Laboratory for Big Data System Computing Technology, Shenzhen University

³Shenzhen Institute of Artificial Intelligence and Robotics for Society

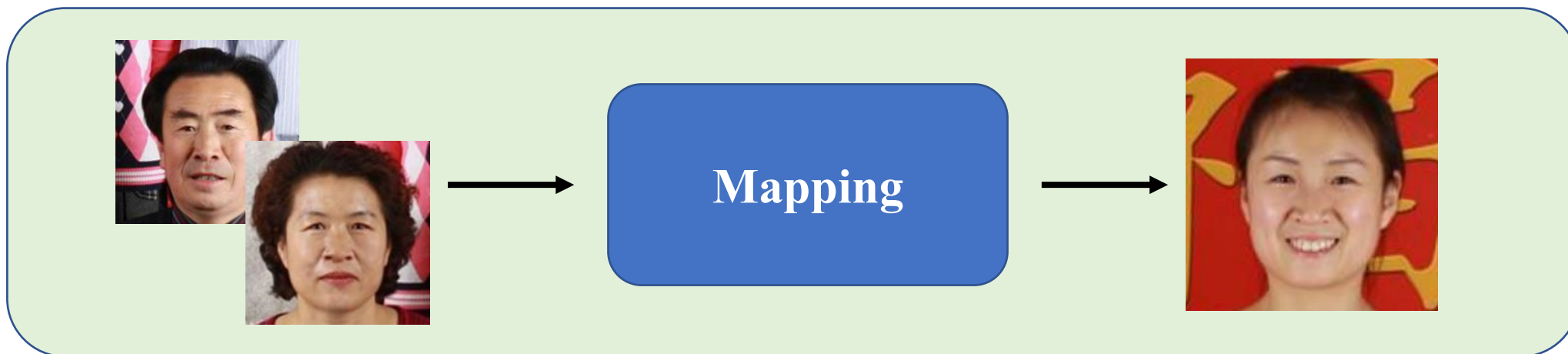
⁴Guangdong Key Laboratory of Intelligent Information Processing, Shenzhen University

⁵School of AI and Advanced Computing, Xi'an Jiaotong-Liverpool University

Github: <https://github.com/CVI-SZU/StyleGene>

Quick Preview

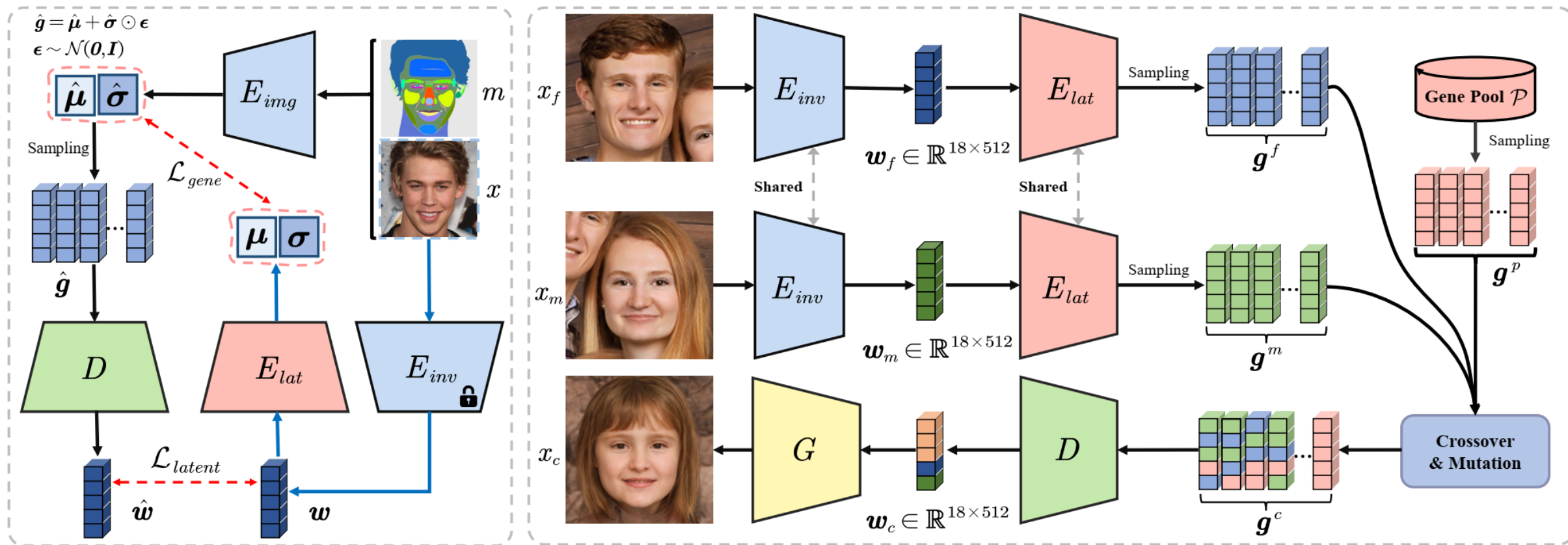
- Synthesizing Descendant Faces using Parental Images



Applications:

- Finding missing children
- Building family trees
- Criminal pursuits
- Social media analysis

Quick Preview



(a) Training Stage

(b) Inference Stage

Motivation

Challenges:

- Missing large-scale, high-quality kinship datasets make visual genetic relations learning difficult.
- Data limitations lead to overfitting in learned genetic relations, resulting in limited diversity among generated descendants.

Objectives:

- Reducing reliance on kinship annotation during the training stage.
- Enhancing facial genetic region controllability and interpretability.
- Improving the diversity of facial mutation traits.
- Minimizing training complexity and overhead.

Contribution

StyleGene Framework:

- Synthesizing high-fidelity kinship faces with controllable facial genetic regions.

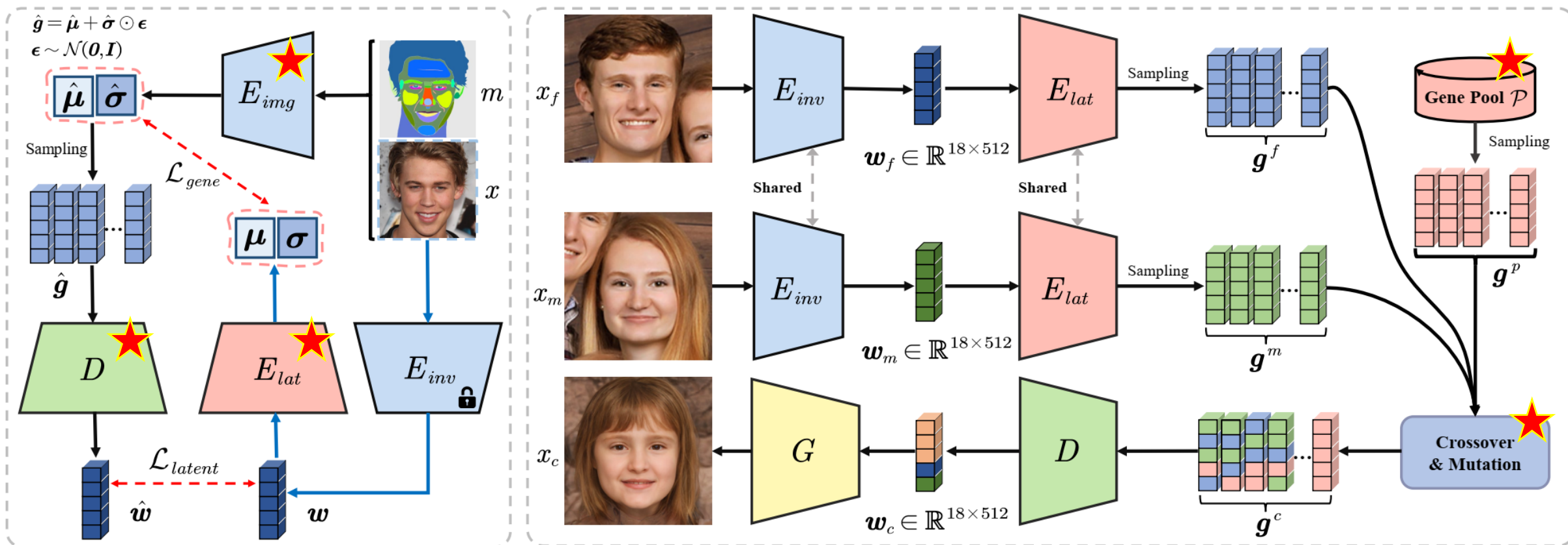
Novel Genetic Strategy:

- Simulating crossover and mutation to generate region-level facial genes (RFGs) for descendants.

Gene Pool Enhancement:

- Increasing kinship face diversity through the Gene Pool concept.

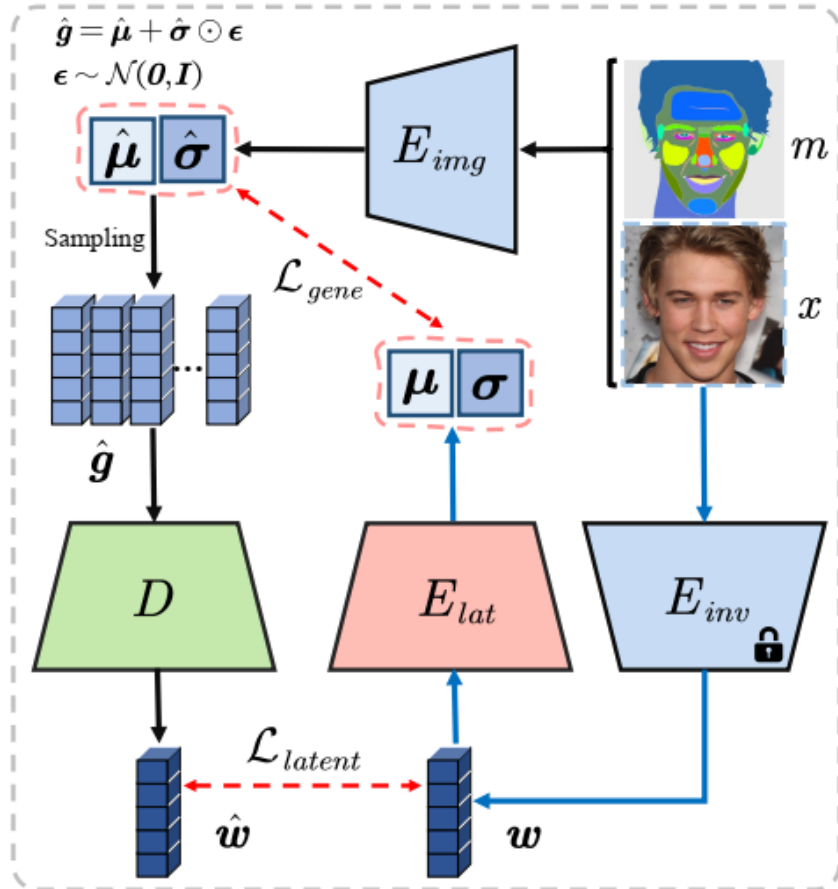
Proposed Method



(a) Training Stage

(b) Inference Stage

Learning Region-level Facial Gene



Reconstruction loss

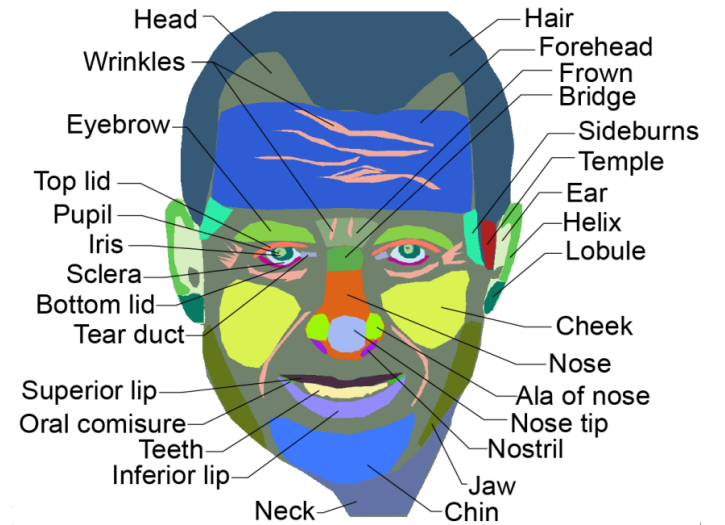
$$\mathcal{L}_{latent} = \| w - \hat{w} \|_2,$$

Decouple latent loss

$$\mathcal{L}_{gene} = \sum_{i=1}^N \left[\| \mu_i - \hat{\mu}_i \|_2 + \| \sigma_i - \hat{\sigma}_i \|_2 \right],$$

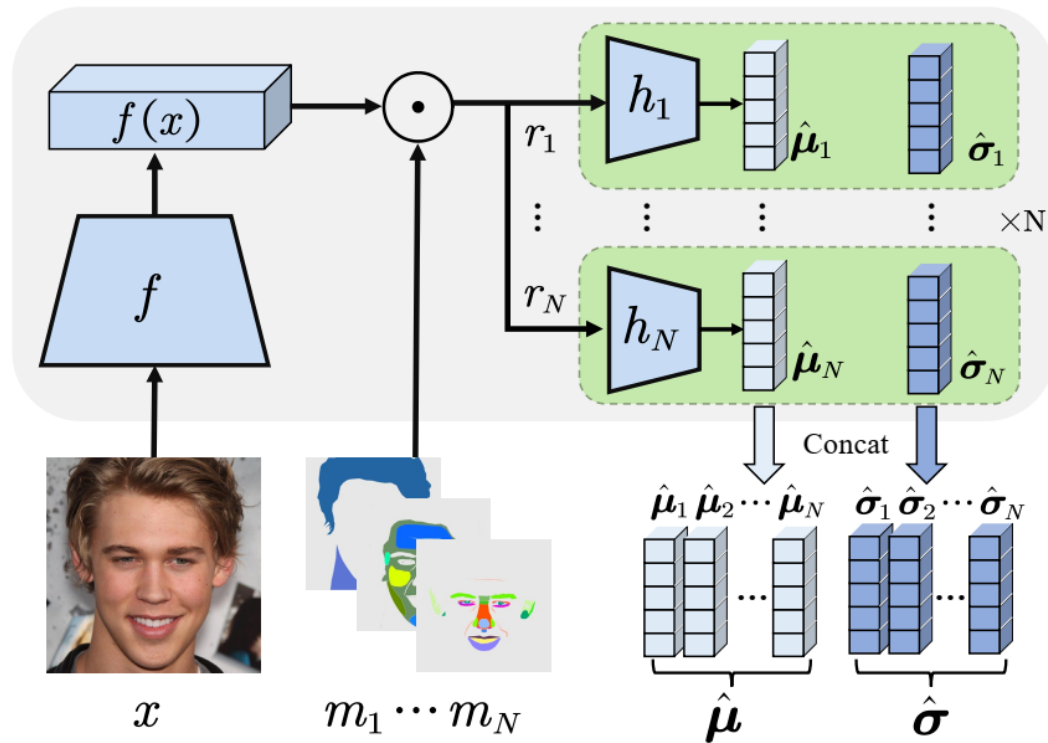
Joint training of all modules

$$\mathcal{L} = \mathcal{L}_{latent} + \lambda \mathcal{L}_{gene},$$



Architecture

Image-based Gene Encoder (IGE) E_{img}



RFG is sampled from a Gaussian distribution:

$$\hat{g}_i \sim q(\hat{g}_i | r_i) = \mathcal{N}(\hat{g}_i; \hat{\mu}_i, \hat{\sigma}_i^2 \mathbf{I}), \quad \hat{g}_i \in \mathbb{R}^{18 \times 512}$$

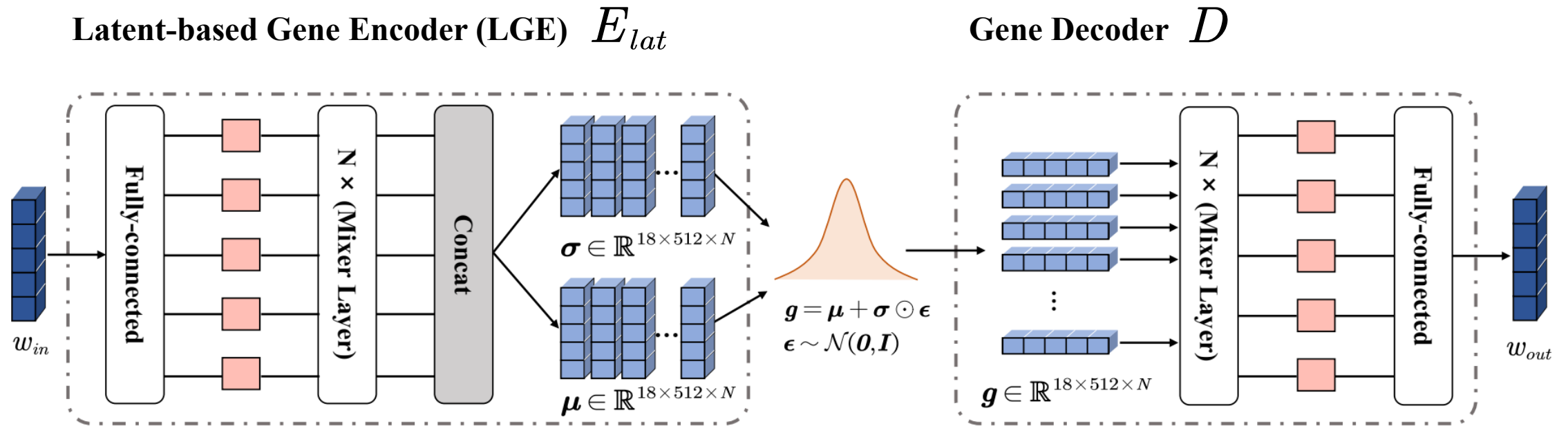
Reparameterization trick:

$$\hat{g}_i = \hat{\mu}_i + \hat{\sigma}_i \odot \epsilon, \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I}),$$

Concatenate all regions RFG:

$$\hat{g} = [\hat{g}_1, \dots, \hat{g}_N] \in \mathbb{R}^{18 \times 512 \times N}$$

Architecture

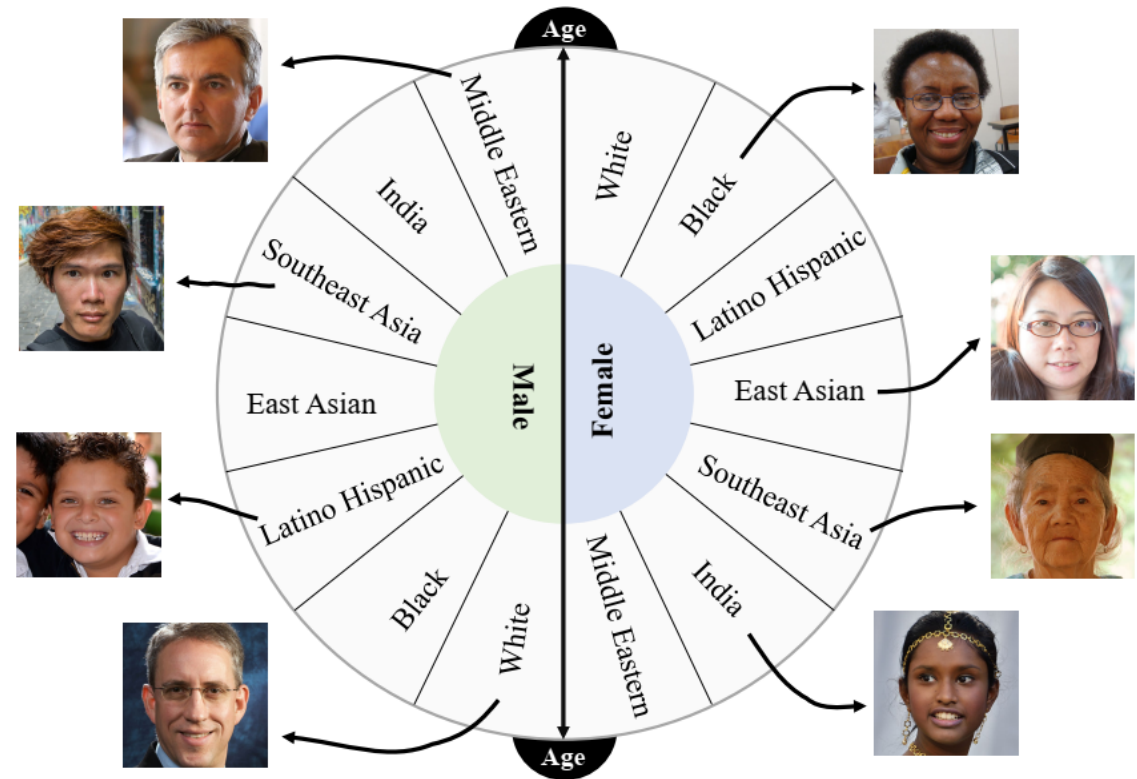


Building a Gene Pool

Step1: Extract RFGs from FFHQ dataset

Step2: Group RFGs by attributes

- Age: 0-2, 3-4, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, and over 70 years old
- Race: White, Black, Indian, East Asian, Southeast Asian, Middle Eastern, and Latino
- Gender: Male and Female



Crossover & Mutation

Given a pair of parental face images, x_f and x_m , we first apply LGE to obtain a set of RFGs $\mathbf{g}^f = [\mathbf{g}_1^f, \dots, \mathbf{g}_N^f]$ and $\mathbf{g}^m = [\mathbf{g}_1^m, \dots, \mathbf{g}_N^m]$.

The RFGs of descendants $\mathbf{g}^c = [\mathbf{g}_1^c, \dots, \mathbf{g}_N^c]$ can be calculated by

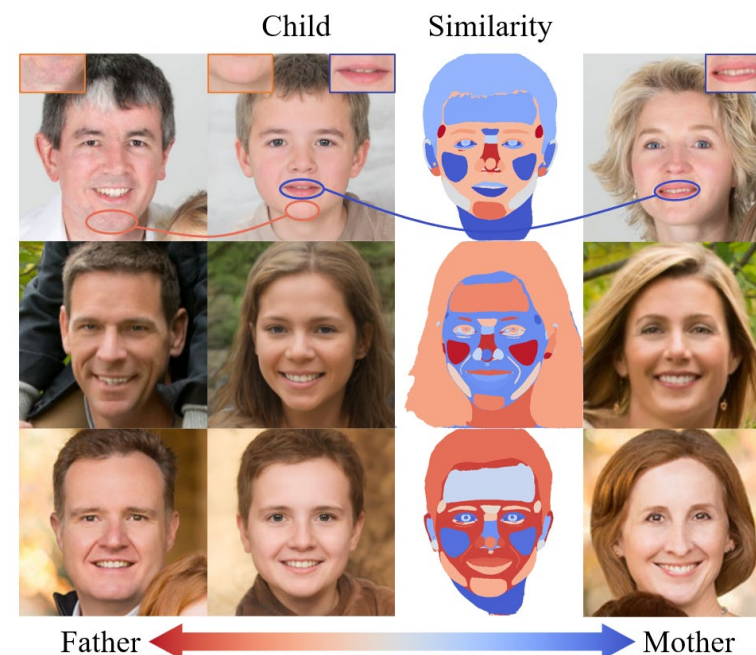
Gene Crossover:

$$\mathbf{g}_i^c = \alpha_i \mathbf{g}_i^f + \beta_i \mathbf{g}_i^m, \quad \alpha_i + \beta_i = 1$$

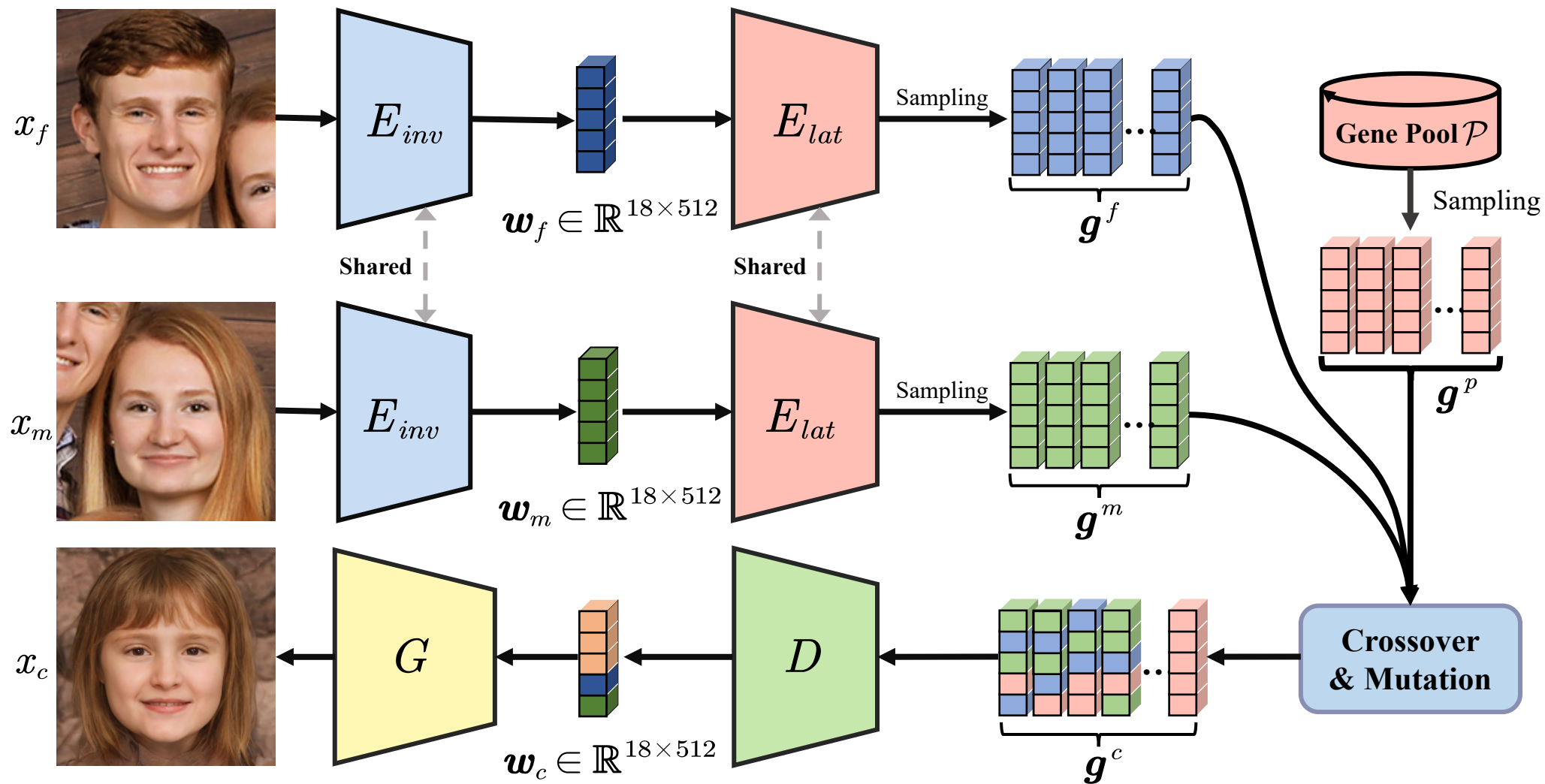
Gene Mutation:

$$\mathbf{g}_i^c = \begin{cases} \mathbf{g}_i^p, & t_i = 1 \\ \alpha_i \mathbf{g}_i^f + \beta_i \mathbf{g}_i^m + \gamma \mathbf{g}_i^p, & t_i = 0 \end{cases},$$

where $\alpha_i + \beta_i = 1 - \gamma$, $\mathbf{g}_i^p = S(\mathcal{P}_c)$, $S(\cdot)$ is the random sampling operator.



Inference Stage



Experiments

Qualitative evaluation

➤ FIW dataset



Father Mother Real Child Ours StyleDNA ChildPredictor ChildGAN DNA-Net

Qualitative evaluation

➤ FF-Database



Father

Mother

Real Child

Ours

StyleDNA

ChildPredictor

Qualitative evaluation

➤ TSKinFace dataset



Father

Mother

Real Children

Ours

StyleDNA

ChildPredictor

CDFS

Quantitative evaluation

➤ Kinship Verification

Table 2. Kinship verification accuracy (%) on the TSKinFace [31], FF-Database [47], and FIW [34] dataset.

Methods	TSKinFace	FF-Database	FIW
StyleDNA [26]	53.15	55.11	49.47
ChildPredictor [47]	58.24	59.62	51.81
StyleGene (Ours)	81.74	80.38	62.29

➤ Diversity Evaluation (LPIPS metric)

Table 3. Quantitative comparison of the diversity of the generated descendants. * means we cropped the face i.e. no hair.

Methods	TSKinFace	FF-Database	FIW
StyleDNA* [26]	0.0756	0.0763	0.0736
ChildPredictor [47]	0.1697	0.1723	0.1750
StyleGene (Ours)*	0.1748	0.1735	0.1740
StyleDNA [26]	0.1559	0.1542	0.1573
StyleGene (Ours)	0.3270	0.3418	0.3279

User Study

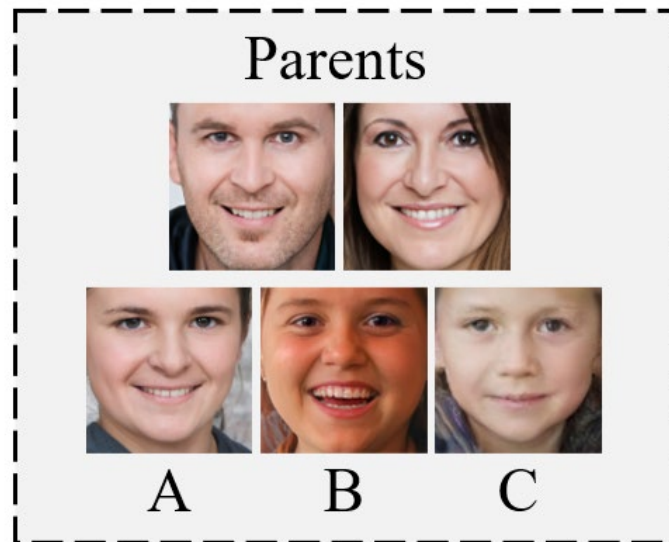


Table 4. The rank of different approaches in user study.

	ChildPredictor	StyleDNA	StyleGene (Ours)
Avg. rank	2.46	2.22	1.32

Distribution of Synthesized Children

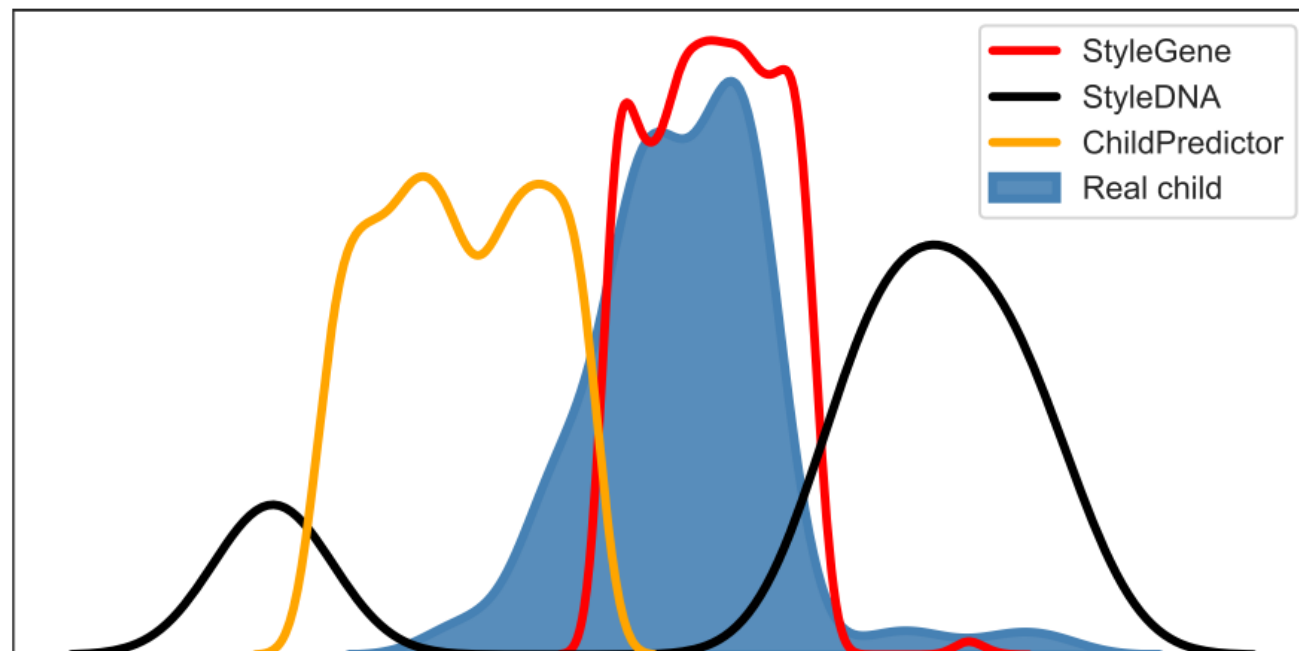


Figure 9. Distribution of real children (blue) and children generated using StyleDNA (black), ChildPredictor (yellow), and ours (red). Best viewed in color.

Conclusion

- StyleGene is a proposed method for synthesizing high-fidelity kinship faces with controllable facial genetic regions.
- A novel genetic strategy is introduced to simulate the crossover and mutation process to generate the facial genes of descendants.
- A Gene Pool is used to increase the diversity of the kinship face during the mutation process.
- Experimental results demonstrate superior realness, similarity with parents, and diversity compared to existing methods.