Attributes Aware Face Generation with Generative Adversarial Networks

Zheng Yuan, Jie Zhang, Shiguang Shan, Xilin Chen
Institute of Computing Technology, Chinese Academy of Sciences, China,
University of Chinese Academy of Sciences, Beijing, China
zheng.yuan@vipl.ict.ac.cn; {zhangjie, sgshan, xlchen}@ict.ac.cn

Project Code

1. Problem

◆ Attribute to Facial Image

blond hair
female
mouth slightly open
arched eyebrows
heavy makeup

2. Related Work and Motivation

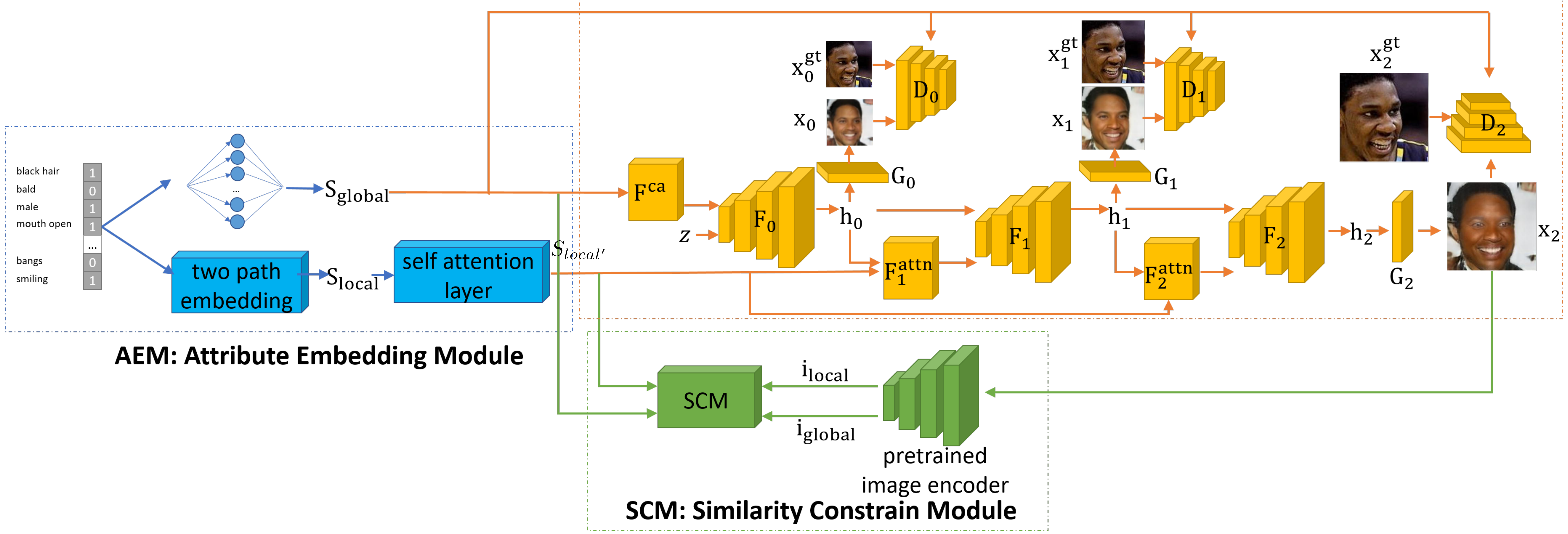
◆ Text to image

➢StackGAN++, AttnGAN, MirrorGAN, etc.
➢the input is different: text vs attribute
➢can not well embed the attribute label

◆ Attribute to image

➢Attribute2sketch2face, Lu et al., Wang et al., etc.
➢the generated images are always low resolution
➢do not consider the relationship between different attributes

3. Overall Structure of AFGAN



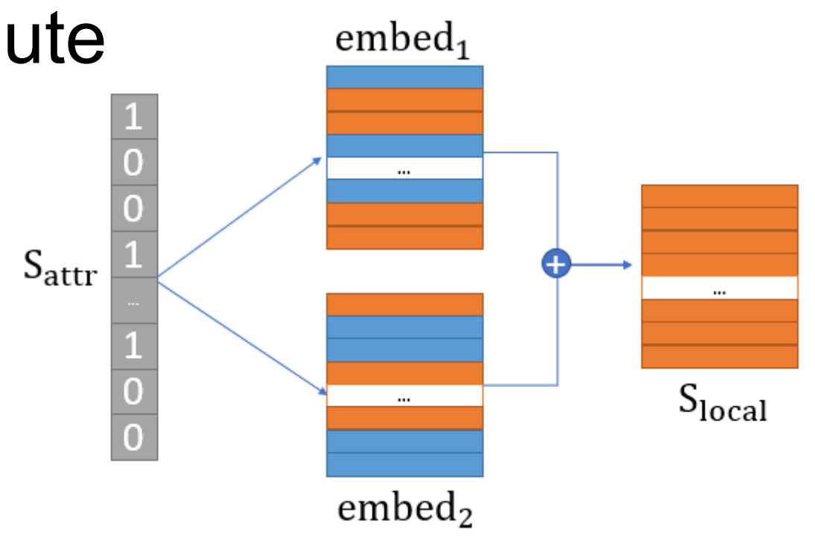
4. Details

◆AEM: Attribute Embedding Module

➢ convert the input face attributes into global and local features respectively

□ two path embedding

➢ well reflect their meanings of the input attribute



$$S_{local} = embed_1 * S_{attr} + embed_2 * (1 - S_{attr})$$

□ self attention layer

➢ model the relationships between different attributes

$$f(x) = W_f * S_{local} \qquad g(x) = W_g * S_{local}$$

$$s_{ij} = f(x_i)^T g(x_j) \qquad \beta_{ij} = \frac{\exp(s_{ij})}{\sum_{i=1}^{N} \exp(s_{ij})}$$

$$h(x) = W_h * S_{local} \qquad S_{local'_j} = \sum_{i=1}^{N} \beta_{ij} h(x_i)$$

$$S_{local'} = (S_{local'_1}, S_{local'_2}, \cdots, S_{local'_N}) \in \mathbb{R}^{C \times N}$$

◆SIGM: Stacked Image Generation Module

➢ gradually generate faces with more details through a three-stage generator

➢ can generate images with high resolution

$$h_0 = F_0(z, F^{ca}(S_{global}))$$

$$h_i = F_i(h_{i-1}, F_i^{attn}(S_{local'}, h_{i-1}))$$

$$x_i = G_i(h_i)$$

◆SCM: Similarity Constrain Module

➢ encode the generated images with a pretrained model: i_{local} and i_{global}

➢ calculate the matching degree between attribute features and image features

➢ the generated images can match the input attributes well

$$s = S_{local'}^T i_{local} \qquad \bar{s}_{ij} = \frac{\exp(s_{ij})}{\sum_{k=1}^{N} \exp(s_{kj})} \quad \alpha_{ij} = \frac{\exp(\gamma_1 \bar{s}_{ij})}{\sum_{k=1}^{2} 89 \exp(\gamma_1 \bar{s}_{ik})}$$

$$c_i = \sum_{j=1}^{289} \alpha_{ij} i_{local_j} \qquad R(c_i, S_{local'}) = \frac{c_i^T S_{local'}}{\|c^i\| \|S_{local'}\|}$$

$$R^{local}(Q, D) = \log(\sum_{i=1}^{N} \exp(\gamma_2 R(c_i, S_{local'})))^{\frac{1}{\gamma_2}}$$

$$R^{global}(Q, D) = \frac{i_{global}^T S_{global}}{\|i_{global}^T\| \|S_{global}\|}$$

5. Objective Function

◆Generator

➢ Overall $\mathcal{L} = \mathcal{L}_G + \mathcal{L}_{SCM}$

➢ In SIGM

$$\mathcal{L}_G = \sum_{i=0}^{2} \mathcal{L}_{G_i} \qquad \mathcal{L}_{G_i} = -\frac{1}{2}\mathbb{E}_{x_i \sim p_{G_i}} [\log (D_i (x_i))] - \frac{1}{2}\mathbb{E}_{x_i \sim p_{G_i}} [\log (D_i (x_i, S_{global}))]$$

➢ In SCM $\mathcal{L}_{SCM} = \mathcal{L}_1^{local} + \mathcal{L}_2^{local} + \mathcal{L}_1^{global} + \mathcal{L}_2^{global}$

$$\mathcal{L}_1^{local} = -\sum_{i=1}^{M} \log P^{local} (D_i | Q_i) \qquad \mathcal{L}_2^{local} = -\sum_{i=1}^{M} \log P^{local} (Q_i | D_i)$$

$$\mathcal{L}_1^{global} = -\sum_{i=1}^{M} \log P^{global} (D_i | Q_i) \qquad \mathcal{L}_2^{global} = -\sum_{i=1}^{M} \log P^{global} (Q_i | D_i)$$

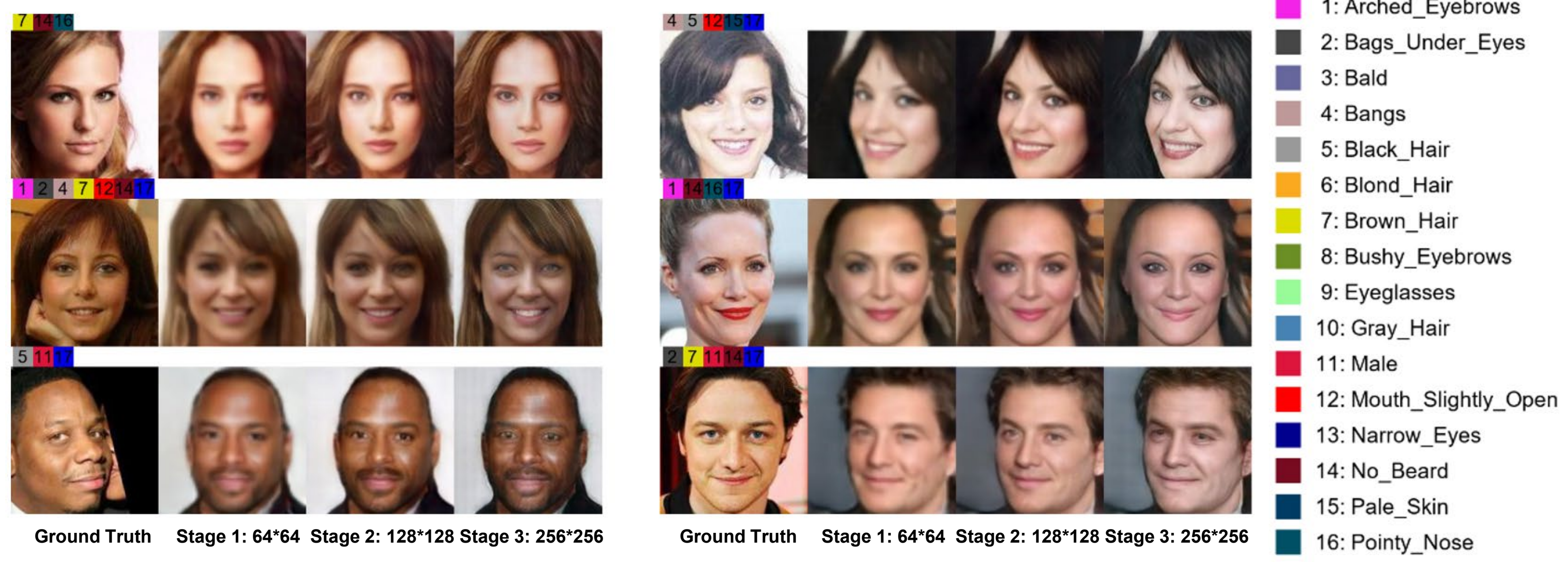
◆Discriminator

$$\mathcal{L}_D = \sum_{i=0}^{2} \mathcal{L}_{D_i}$$

$$\begin{aligned}\mathcal{L}_{D_i} = &- \frac{1}{2}\mathbb{E}_{x_i^{gt} \sim p_{data_i}} \left[\log D_i \left(x_i^{gt}\right)\right] - \frac{1}{2}\mathbb{E}_{x_i \sim p_{G_i}} [\log (1 - D_i (x_i))] \\ &- \frac{1}{2}\mathbb{E}_{x_i^{gt} \sim p_{data_i}} \left[\log D_i \left(x_i^{gt}, S_{global}\right)\right] \\ &- \frac{1}{2}\mathbb{E}_{x_i \sim p_{G_i}} [\log (1 - D_i (x_i, S_{global}))]\end{aligned}$$

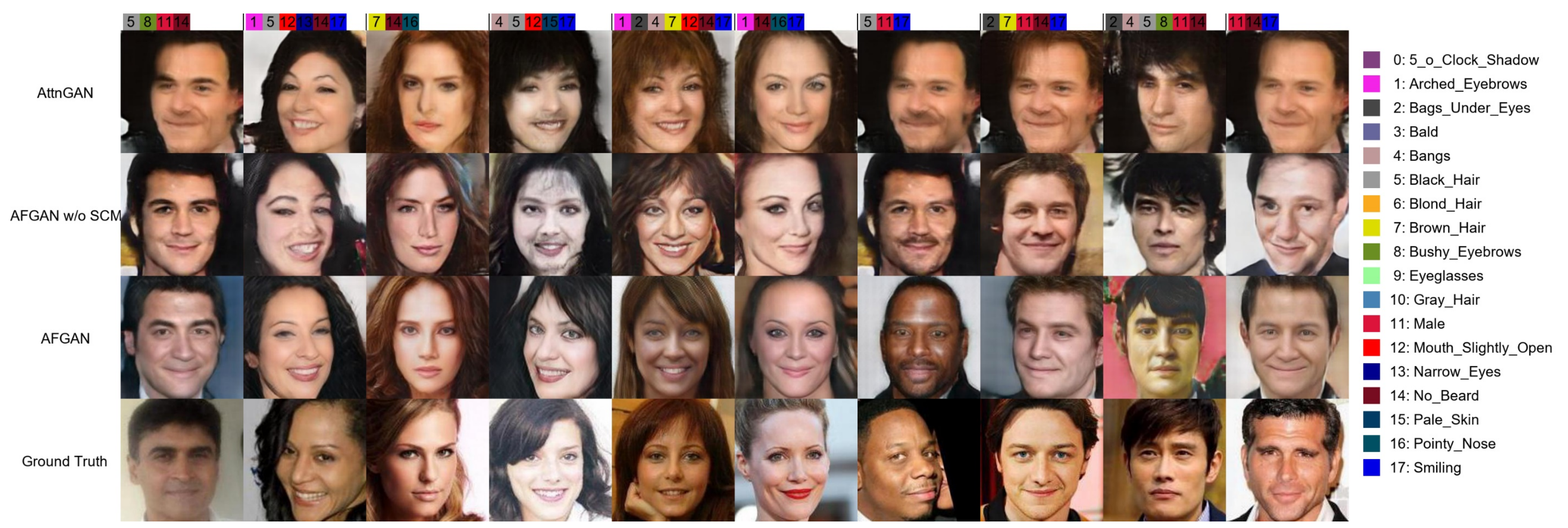
6. Experimental Results

◆The generated face images of three stages in SIGM



◆The comparison of generated images with other methods

□ Qualitive results



□ Quantitative results

	BRISQUE↓	IS↑	FID↓	MS-SSIM↓
AttnGAN	62.843	5.124	40.254	0.398
Wang et al.	——	2.2	43.8	——
AFGAN(ours)	**35.979**	**5.853**	**36.607**	**0.347**

Setting	Classification accuracy
AttnGAN	0.902
AFGAN w/o AEM	0.924
AFGAN w/o SCM	0.940
AFGAN(ours)	**0.955**

ICPR2020@ Milan, Italy