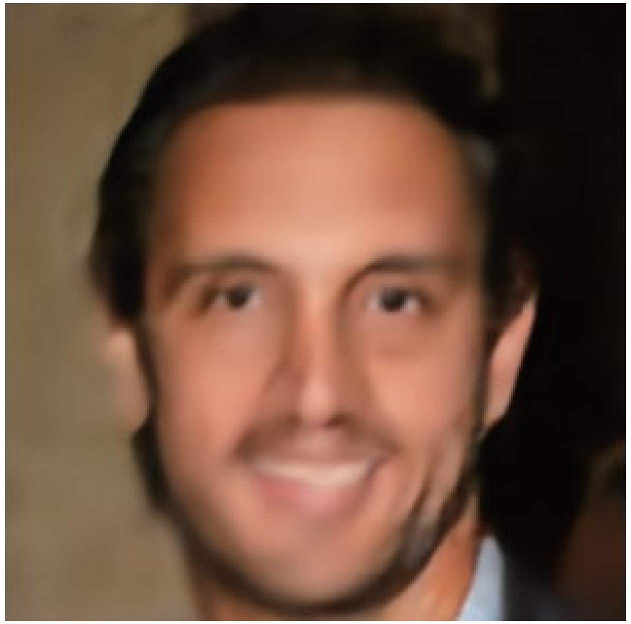
SUSTech Southern University of Science and Technology

Analyzing and Combating Attribute Bias for Face Restoration

Zelin Li, Dan Zeng , Xiao Yan, Qiaomu Shen, Bo Tang

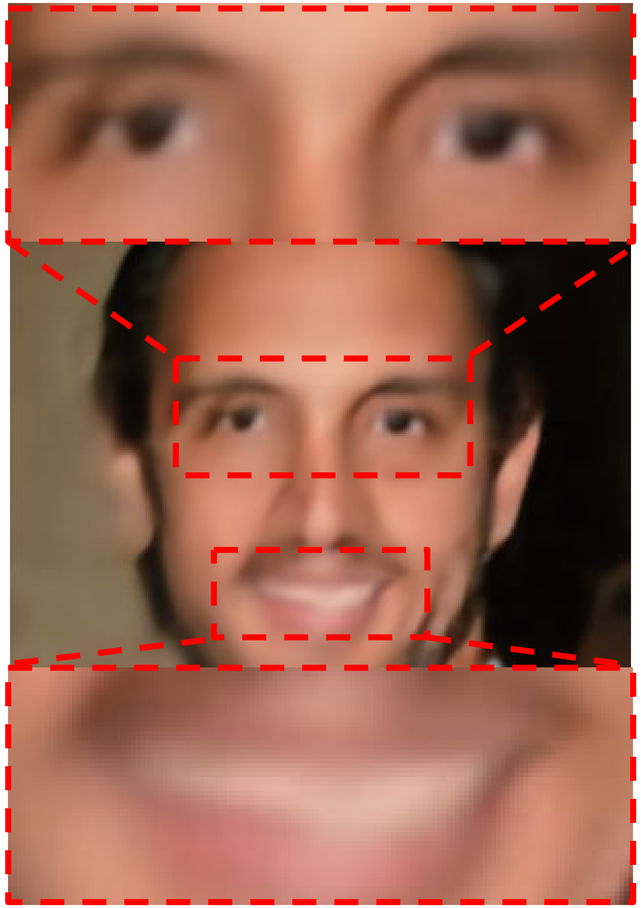
Background

- ➢Face restoration: Recover high-quality (HQ) faces from low-quality (LQ) faces
  - ■ Super-resolution, Denoise, Deblur, etc
- ➢Previous Problem: Over-smooth



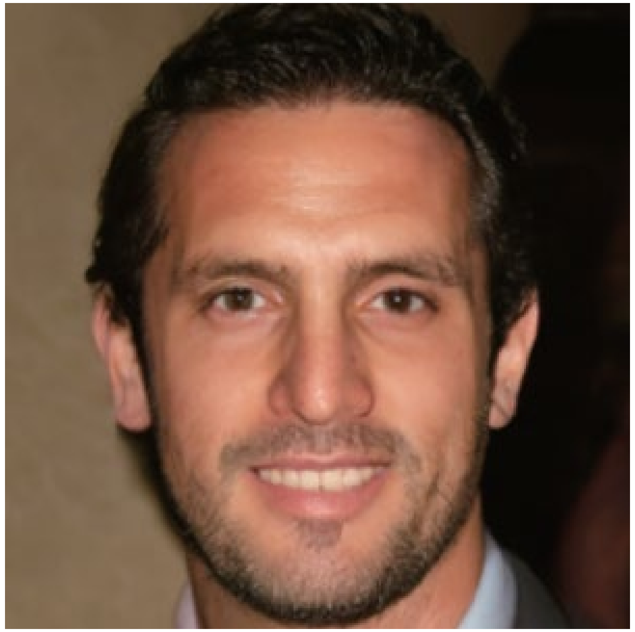
Background

- ➢Face restoration: Recover high-quality (HQ) faces from low-quality (LQ) faces
  - ■ Super-resolution, Denoise, Deblur, etc
- ➢Previous Problem: Over-smooth



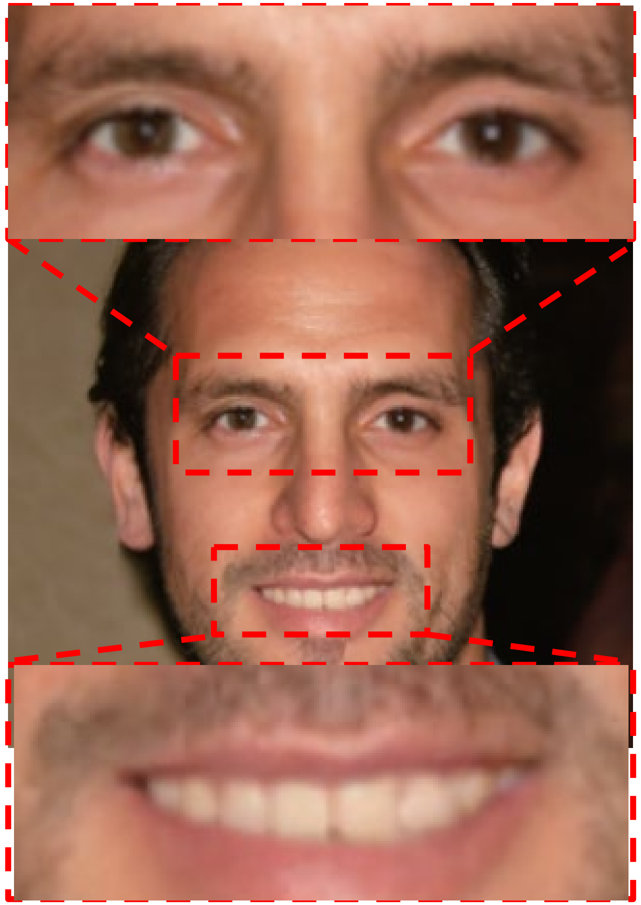
Background

- ➢Face restoration: Recover high-quality (HQ) faces from low-quality (LQ) faces
  - ■ Super-resolution, Denoise, Deblur, etc
- ➢Leveraging generative prior



Background

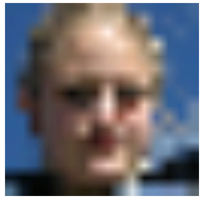
- ➢Face restoration: Recover high-quality (HQ) faces from low-quality (LQ) faces
  - ■ Super-resolution, Denoise, Deblur, etc
- ➢Leveraging generative prior



Attribute Bias

- ➢Key face attributes are dramatically different

Input



16×16

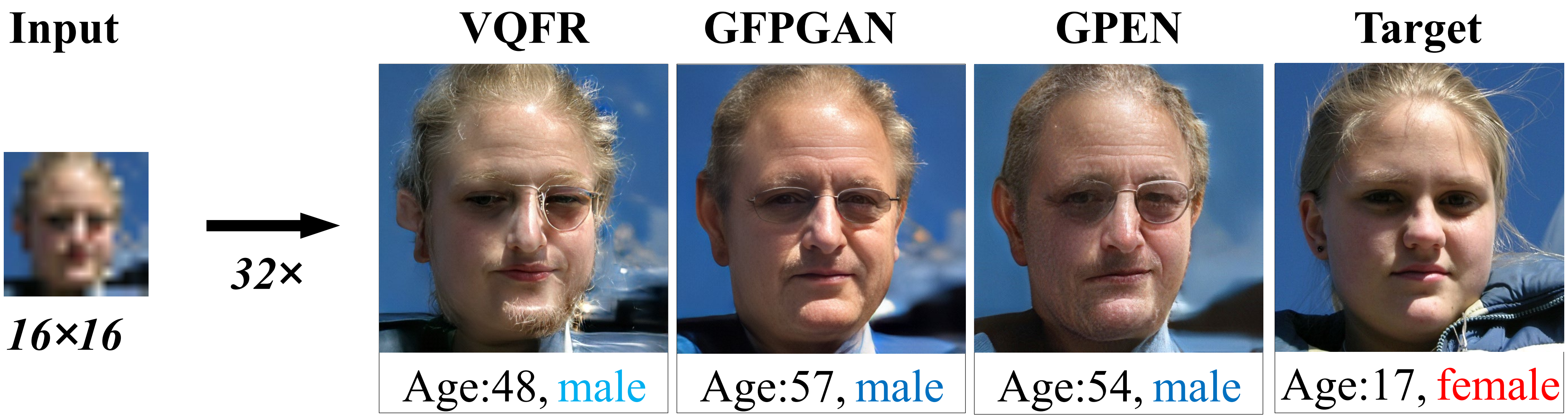
32×

Target

Age:17, female

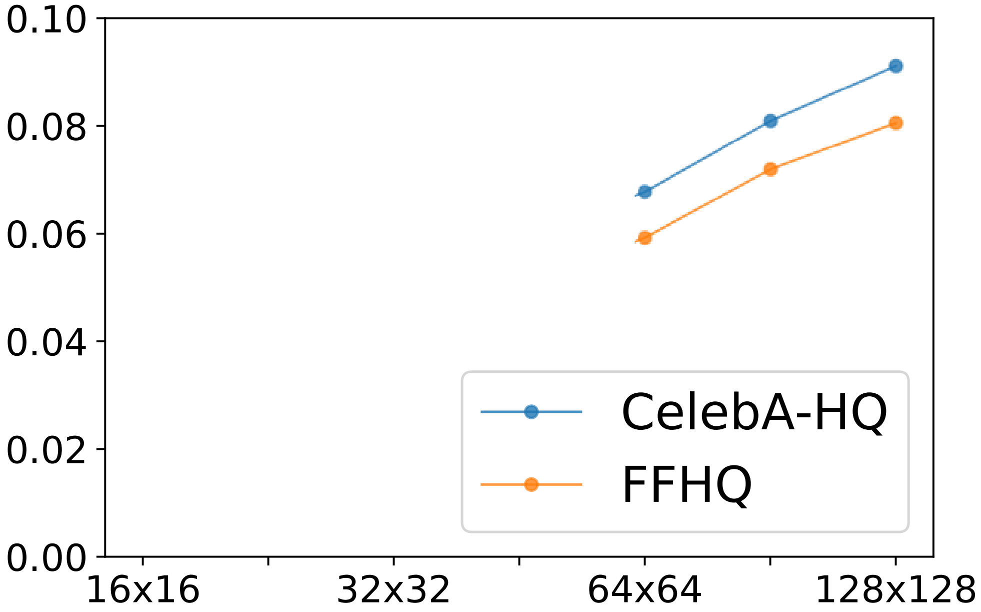
Attribute Bias

- ➢Key face attributes are dramatically different

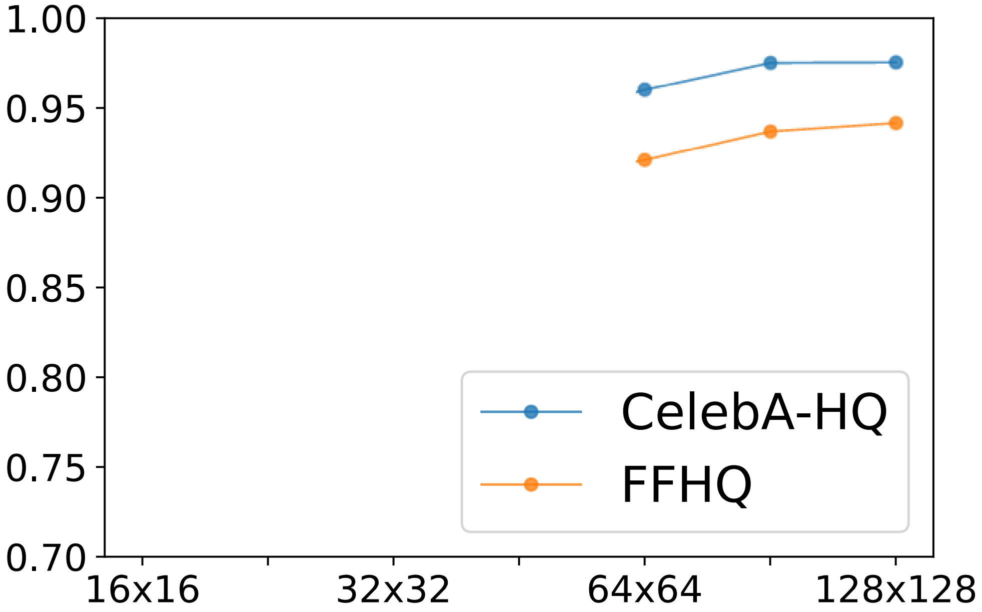


Analysis on Attribute Bias

- ➢Attribute information loses as image resolution decreases



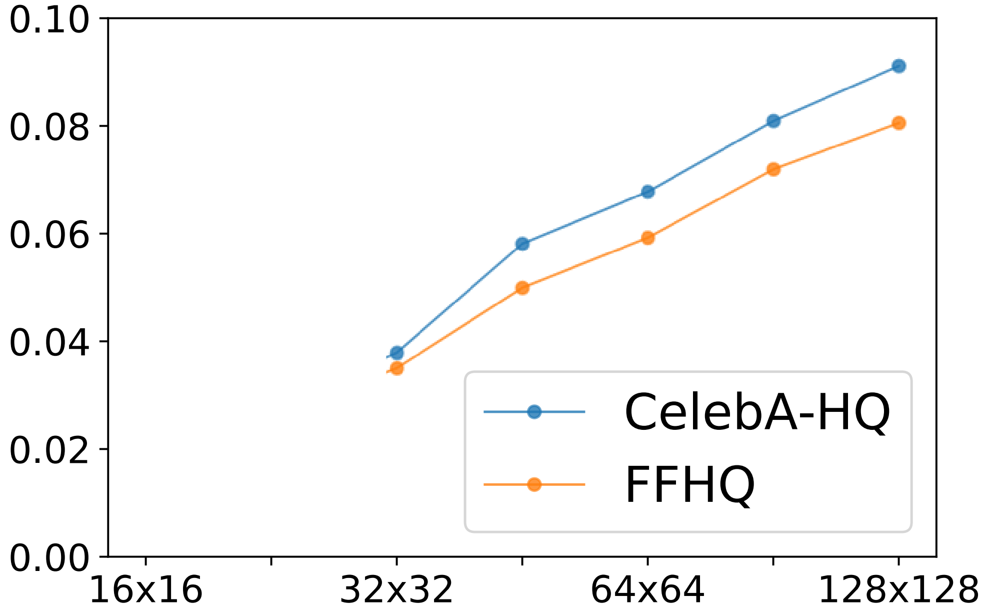
(a) Age Confidence.



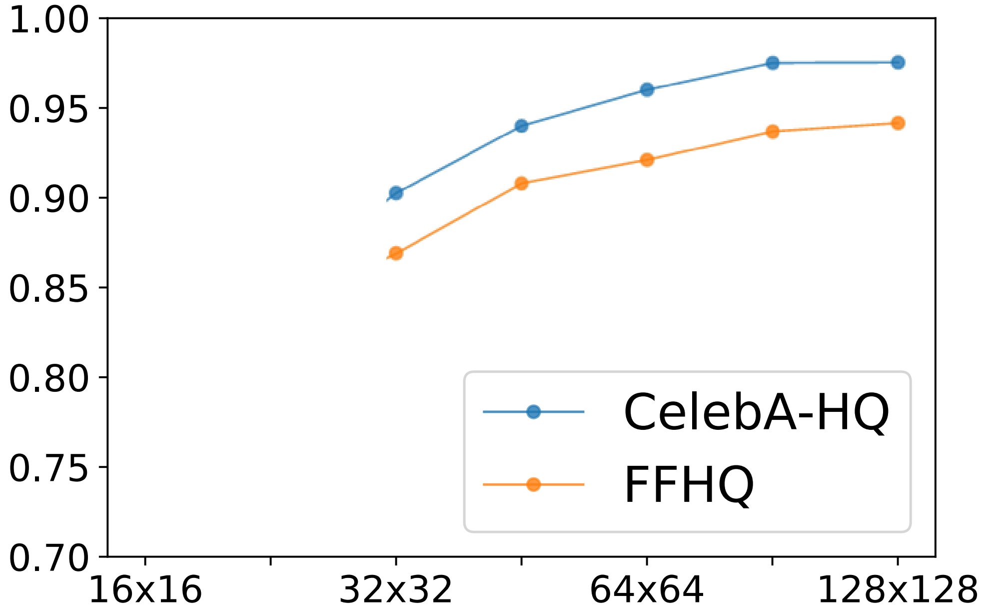
(b) Gender Confidence.

Analysis on Attribute Bias

➢Attribute information loses as image resolution decreases



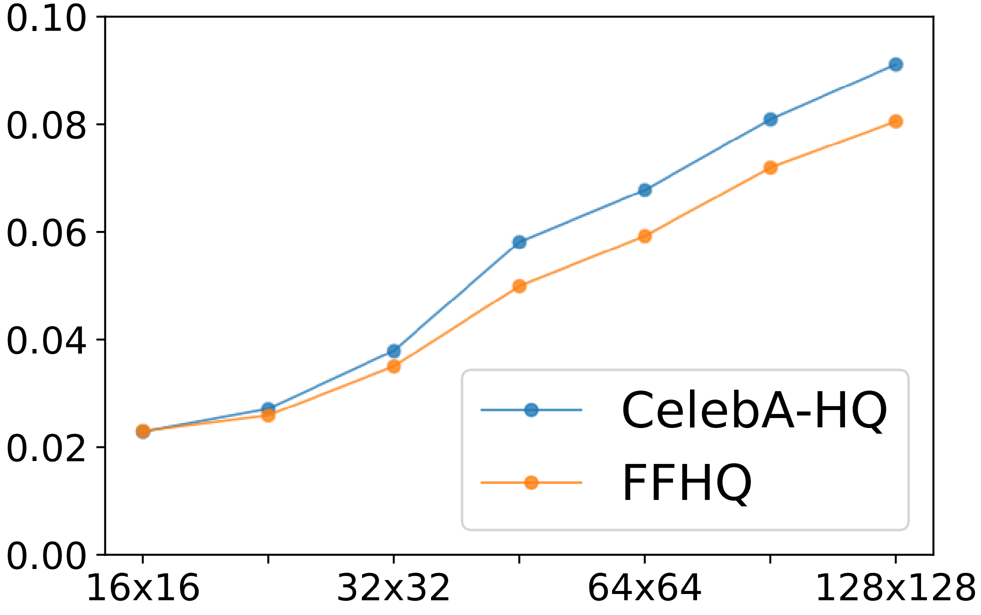
(a) Age Confidence.



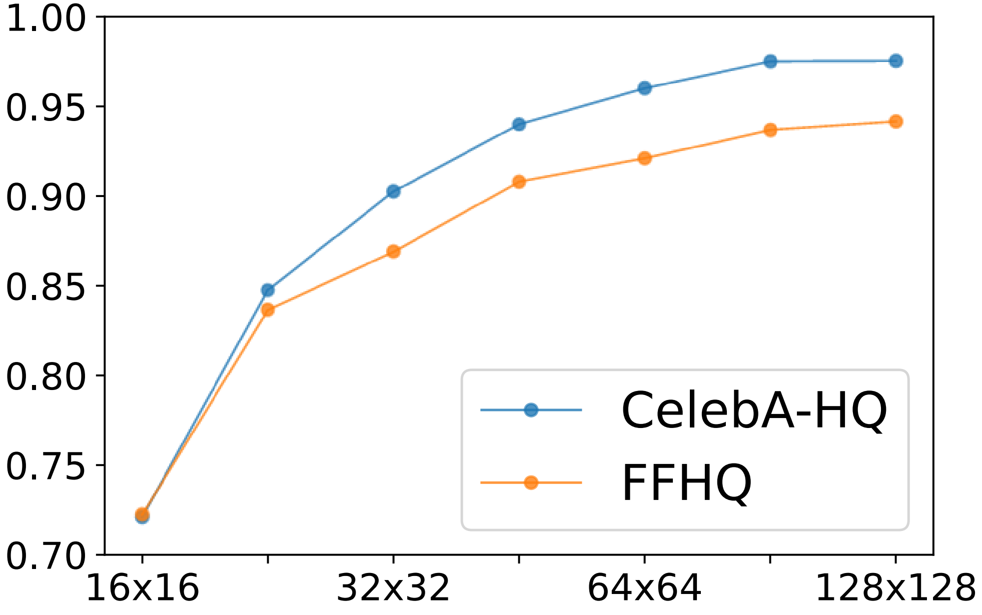
(b) Gender Confidence.

Analysis on Attribute Bias

- Attribute information loses as image resolution decreases

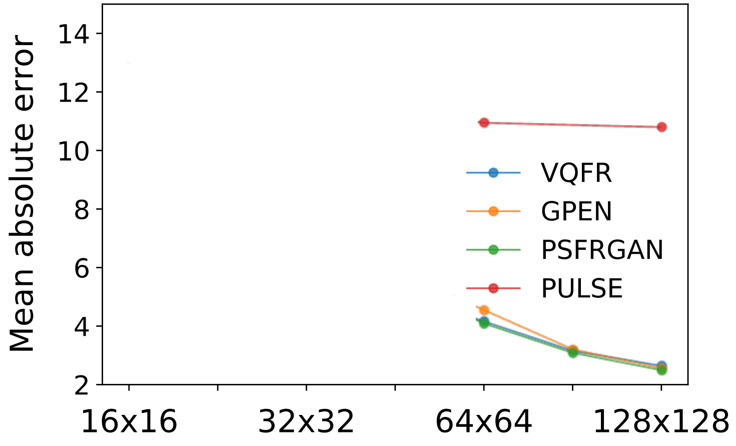


(a) Age Confidence.

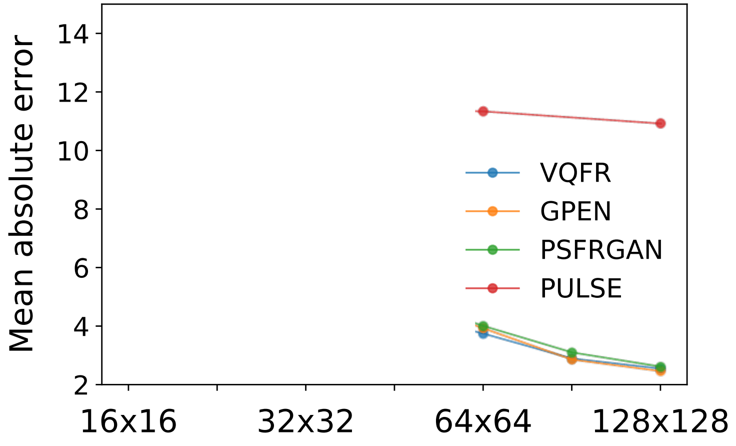


(b) Gender Confidence.

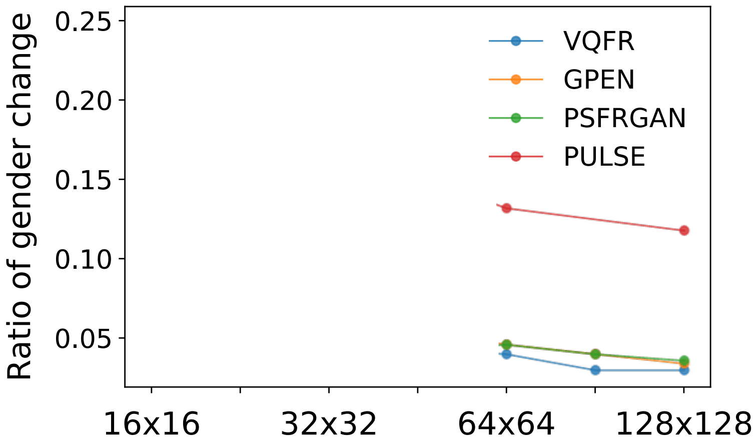
Analysis on Attribute Bias



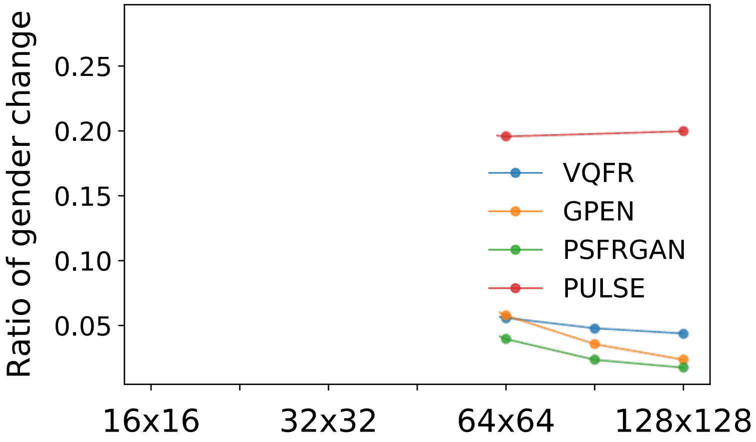
(a) CelebA-HQ.



(b) FFHQ.



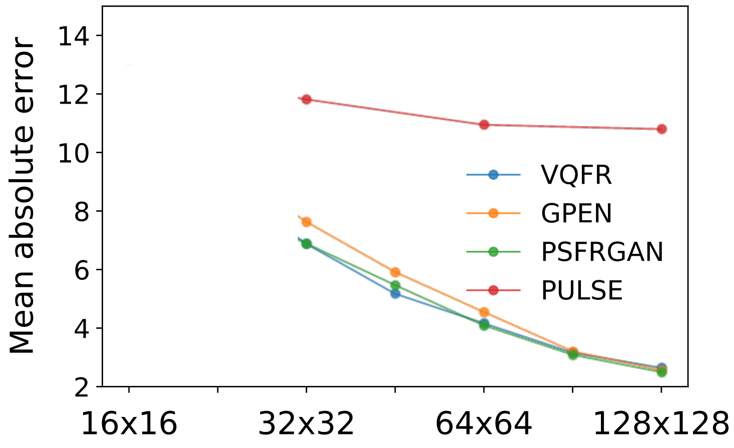
(c) CelebA-HQ.



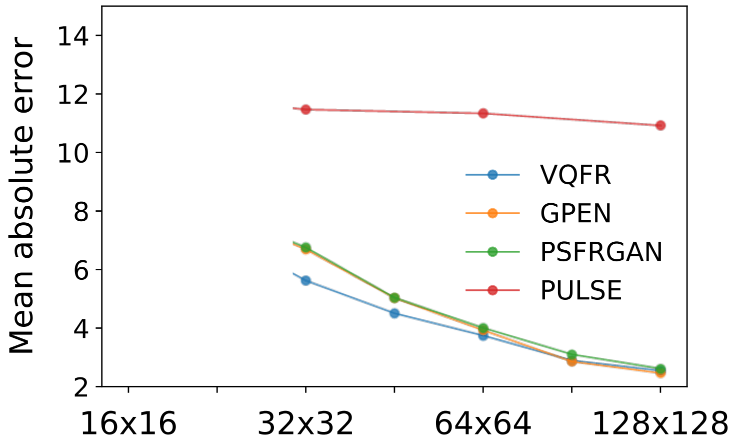
(d) FFHQ.

- Attribute bias enlarges as image resolution decreases

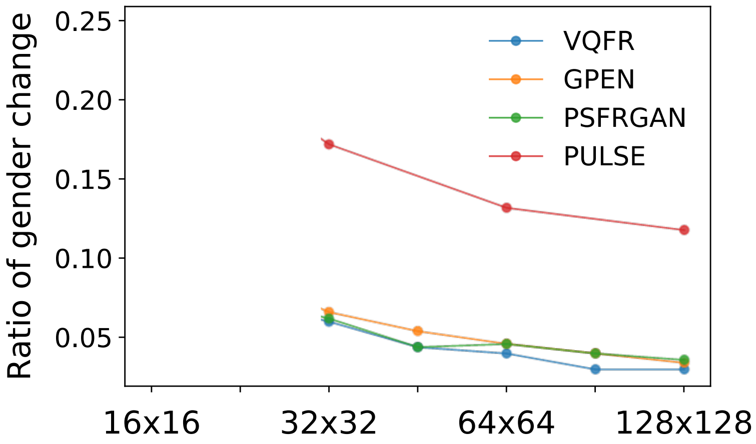
Analysis on Attribute Bias



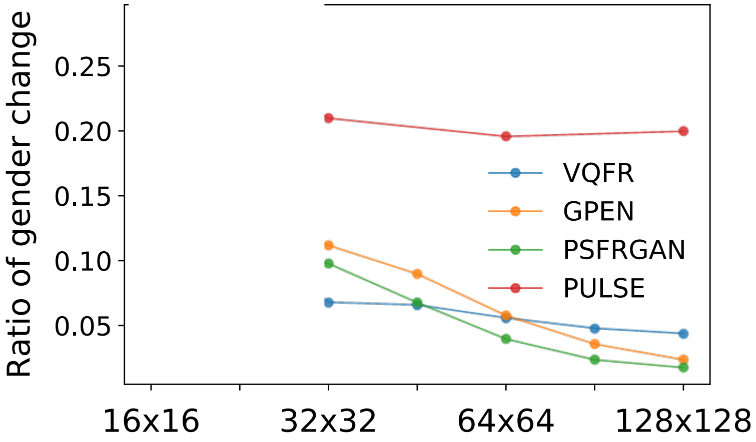
(a) CelebA-HQ.



(b) FFHQ.



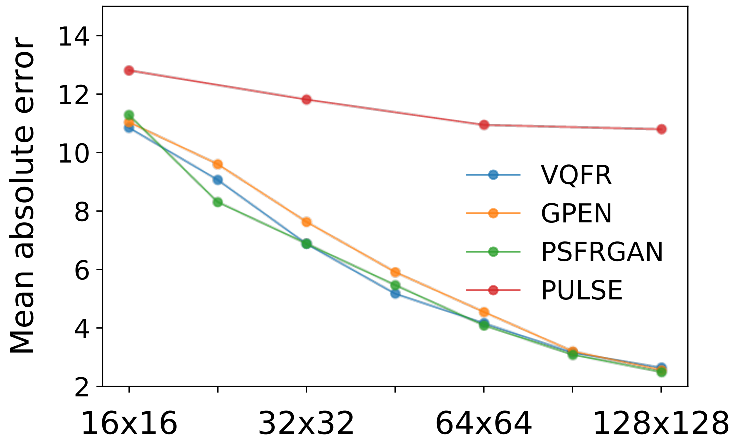
(c) CelebA-HQ.



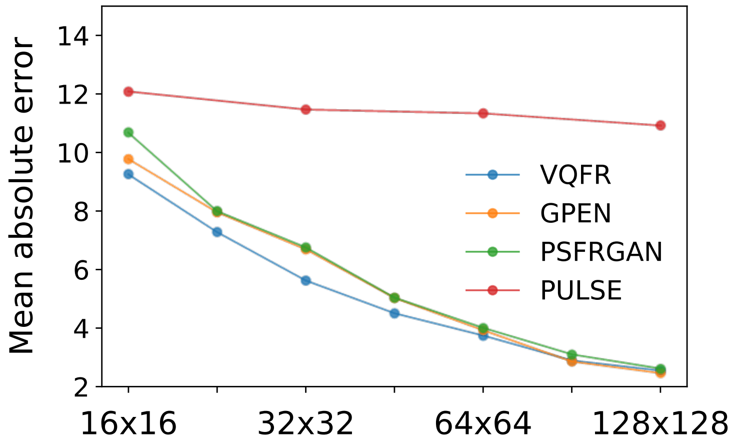
(d) FFHQ.

- Attribute bias enlarges as image resolution decreases

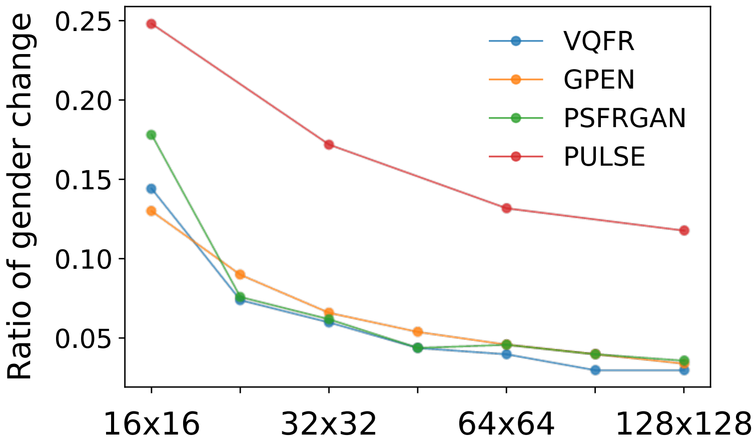
Analysis on Attribute Bias



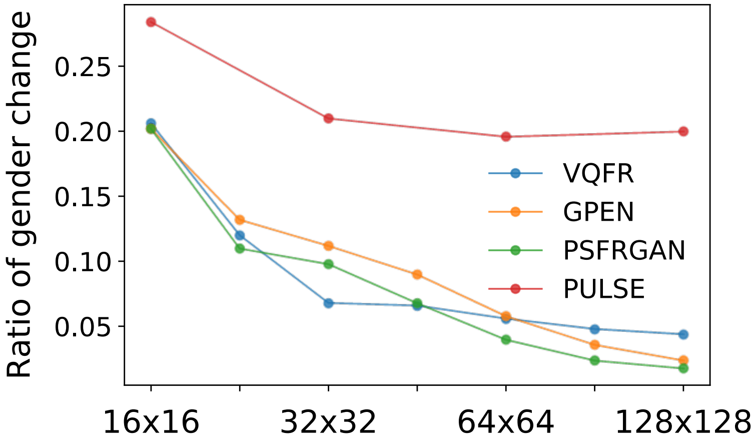
(a) CelebA-HQ.



(b) FFHQ.



(c) CelebA-HQ.

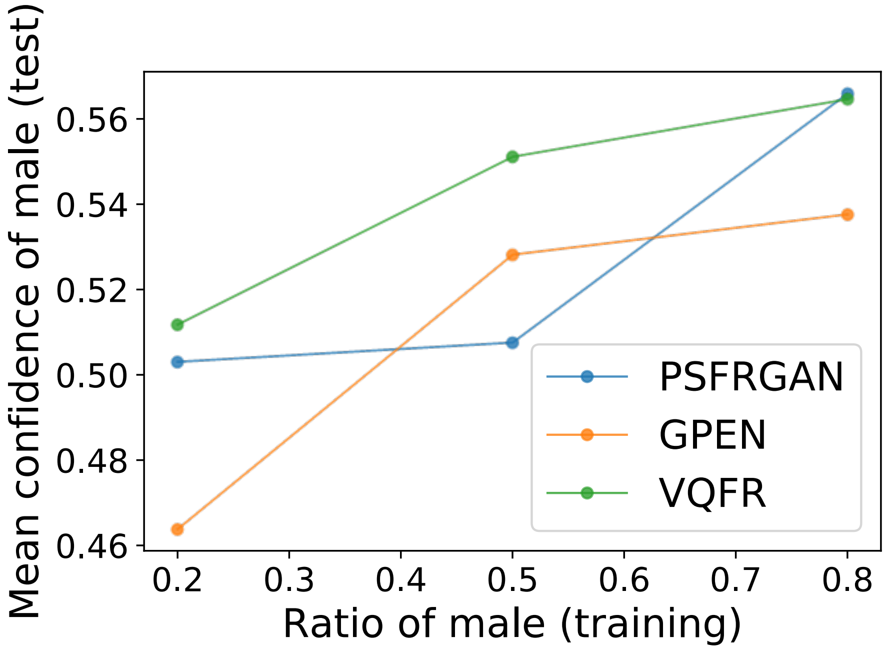


(d) FFHQ.

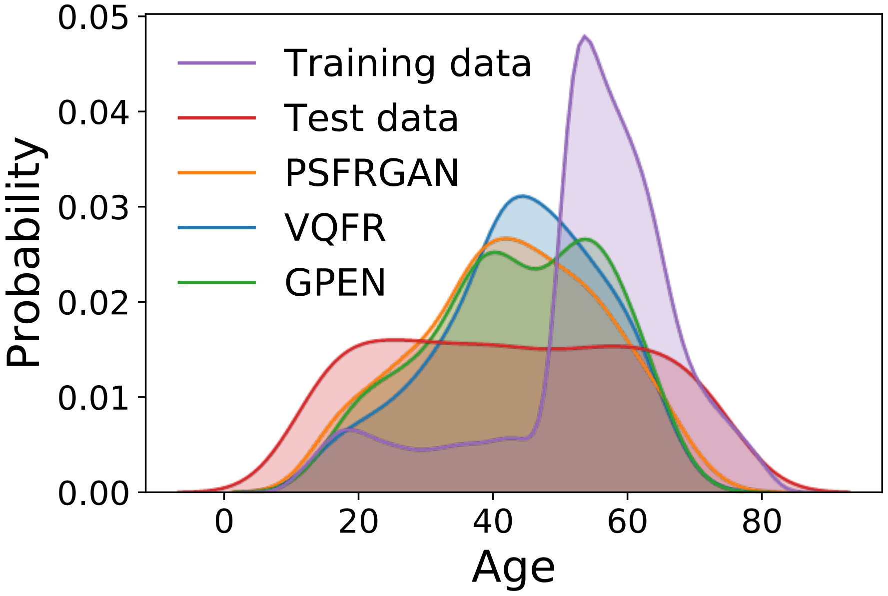
- ➢ Attribute bias enlarges as image resolution decreases

Analysis on Attribute Bias

➢ Affected by the training data distribution



(a)



(b)

Challenges

- The causes are inevitable
  - Lack of attribute information
  - Training data prior

Challenges

- The causes are inevitable
  - Lack of attribute information – Degradation is unavoidable
  - Training data prior – Hard to collect a large dataset with balanced attribute distribution

Challenges

- The causes are inevitable
  - Lack of attribute information – Degradation is unavoidable
  - Training data prior – Hard to collect a large dataset with balanced attribute distribution
- Attribute information is obtainable
  - Witness description, Actor profile

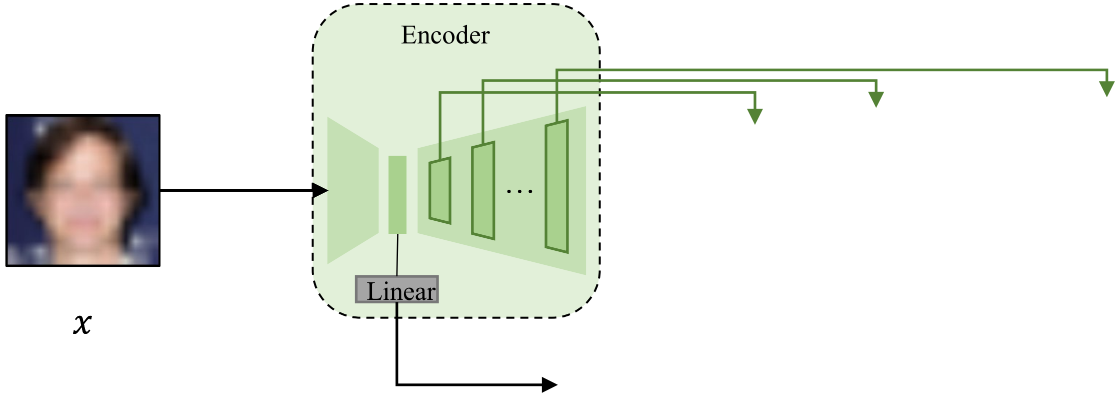
Challenges

- The causes are inevitable
  - Lack of attribute information – Degradation is unavoidable
  - Training data prior – Hard to collect a large dataset with balanced attribute distribution
- Attribute information is obtainable
  - Witness description, Actor profile
- Given attribute information, do image restoration
  - Align attribute & image information
  - Let the attribute have a fine-grained impact on the restoration

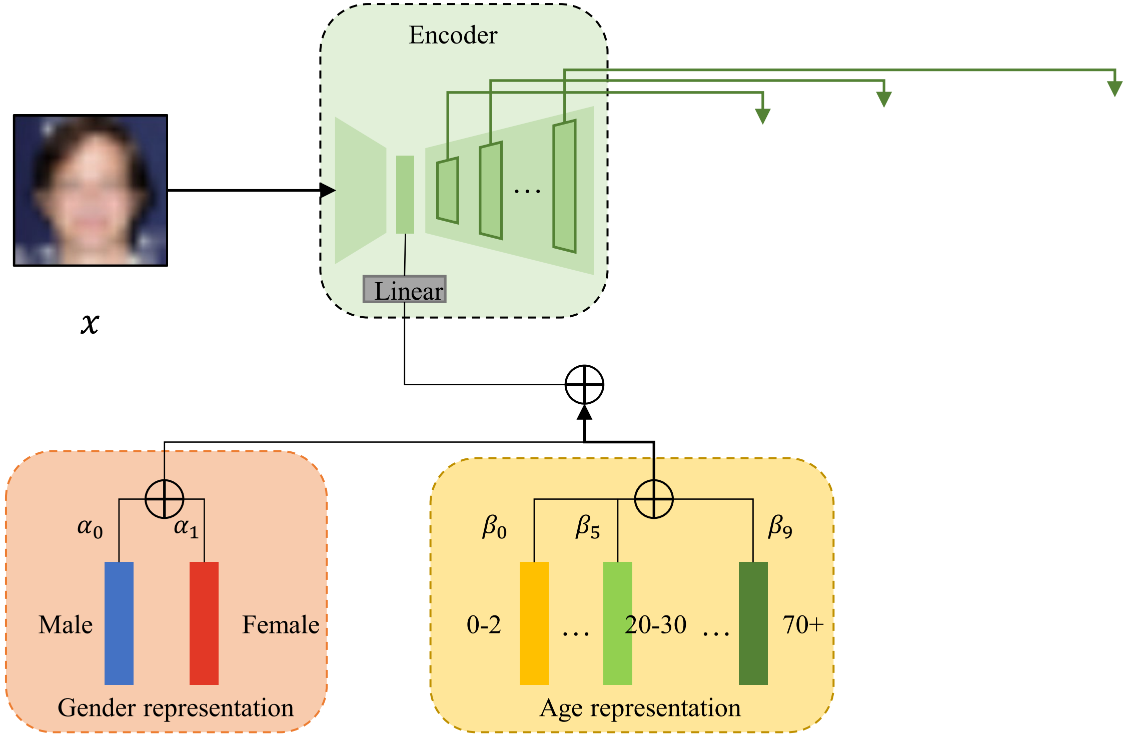
Methodology

x

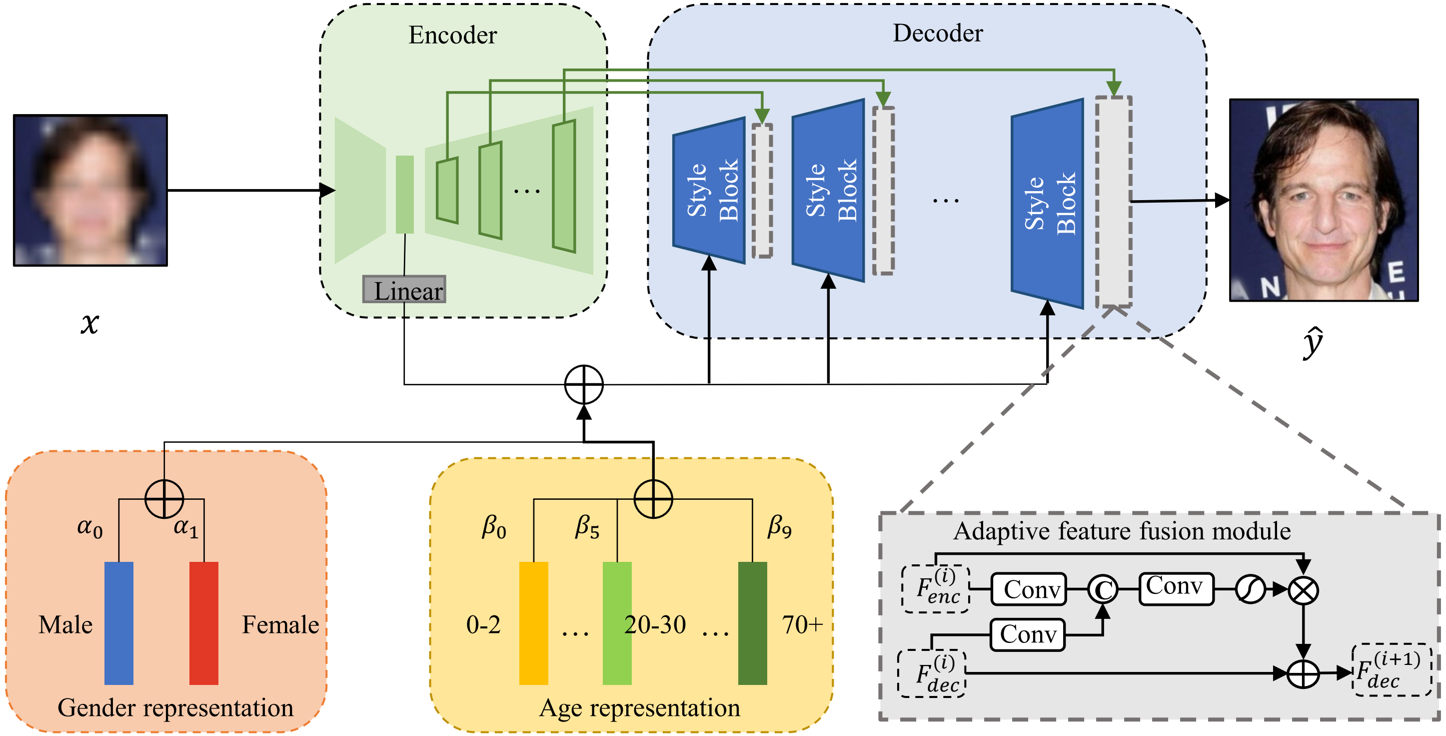
Methodology



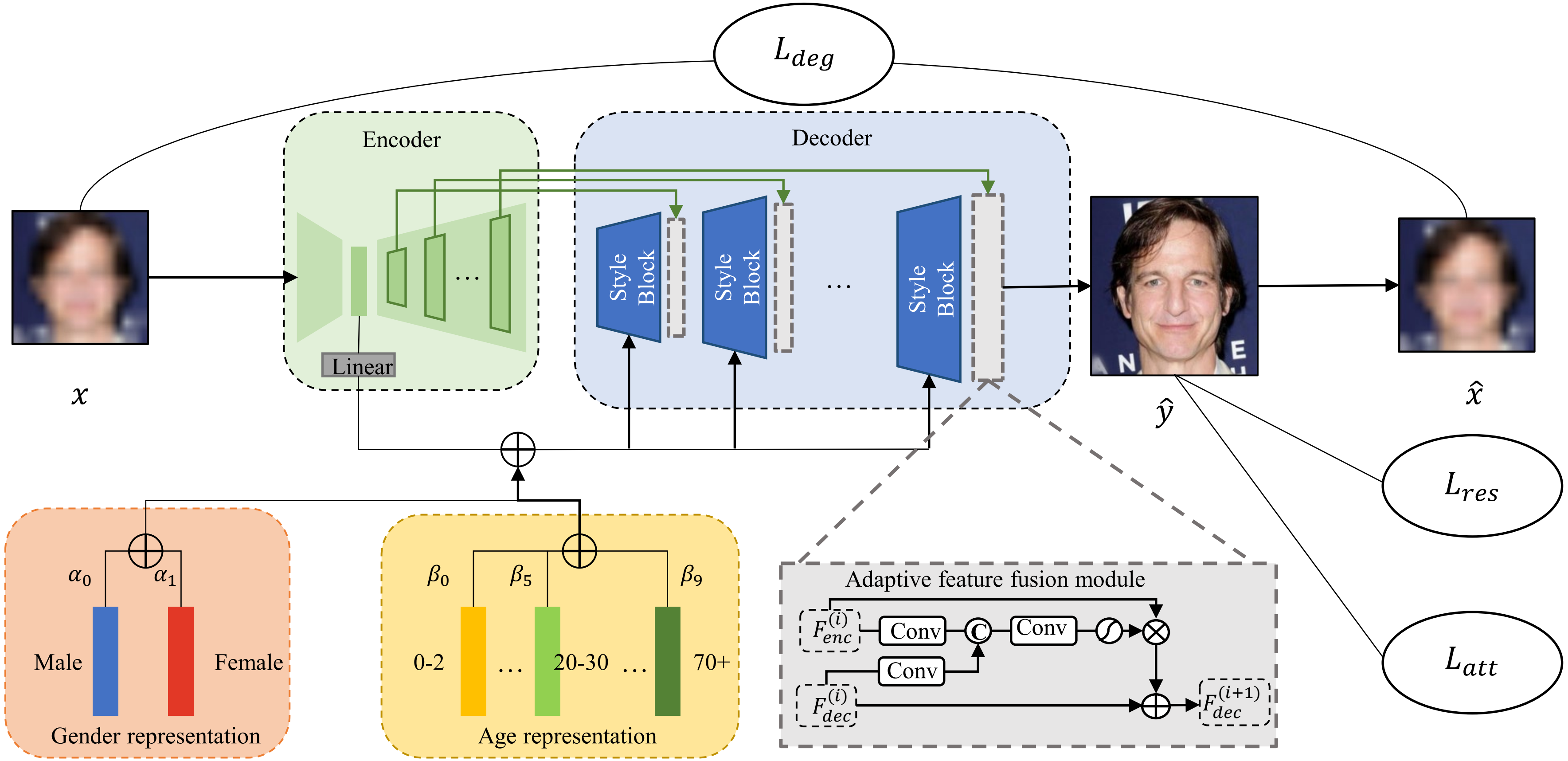
Methodology



Methodology



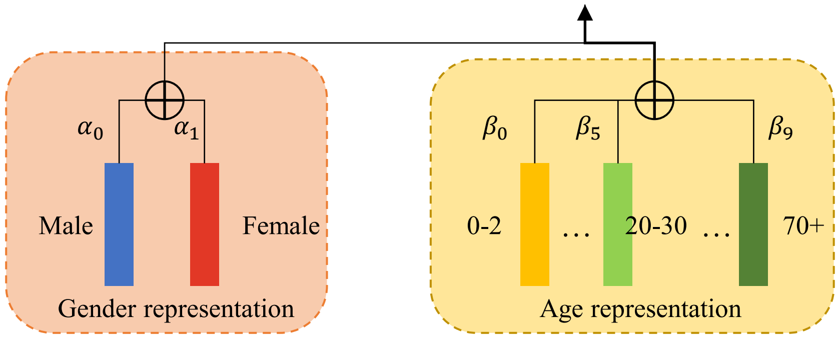
Methodology



Attribute Representation

- Methodology:
  - Determine the attribute weights
  - Base latent vector + weighted sum of attribute representations
  - Update the attribute representations through back-propagation (in the training phase)

- Benefits:
  - Fine-grained control by weights
  - Cheap to extend to more attributes



Two-objective Optimization

➢ Image Restoration

◆ Reconstruction Loss

$$\mathcal{L}_{rec} = \lambda_{L_{pix}} \|\hat{y} - y\|_1 + \mathcal{L}_{per}$$
$$\mathcal{L}_{per} = \|\phi(\hat{y}) - \phi(y)\|_1 + \lambda_{style} \|\mathrm{Gram}(\phi(\hat{y})) - \mathrm{Gram}(\phi(y))\|_1$$

◆ Adversarial Loss

$$\mathcal{L}_{adv,D} = \mathbb{E}_{\hat{y}}[\mathrm{Softplus}(D(\hat{y}))] + \mathbb{E}_{y}[\mathrm{Softplus}(-D(y))]$$
$$\mathcal{L}_{adv,G} = \mathbb{E}_{\hat{y}}[\mathrm{Softplus}(-D(\hat{y}))]$$

➢ Attribute Consistency

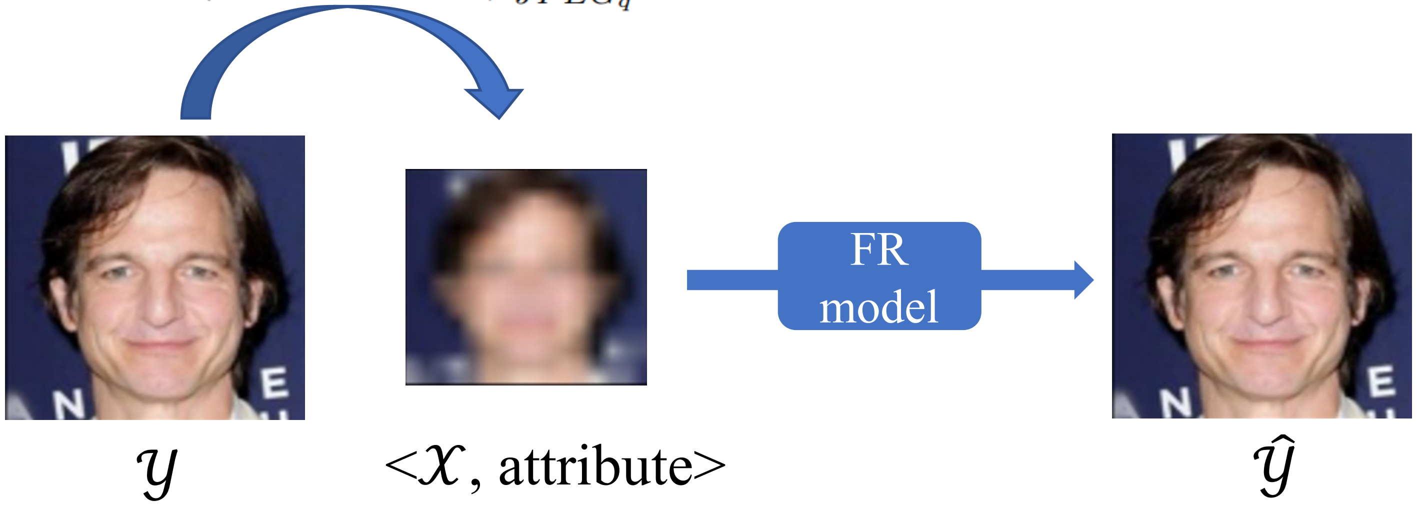
◆ Attribute Loss

$$\mathcal{L}_{att} = \mathrm{CE}(a, P(a|\hat{y}))$$

ϕ : Pretrained VGG-16
D : Discriminator
Gram : Gram matrix
$\hat{y}$: Restored image
y : Target image

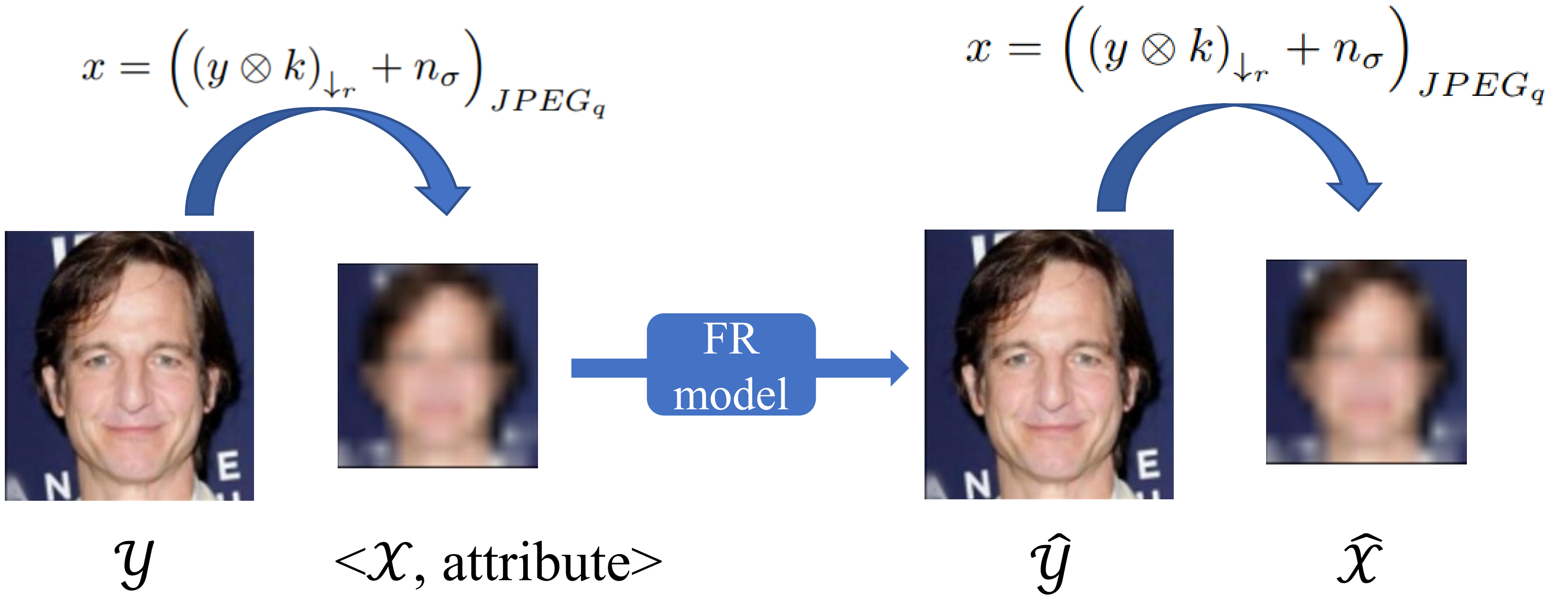
Training Strategy

$$x = \left((y \otimes k)_{\downarrow_r} + n_\sigma \right)_{JPEG_q}$$



- For loss calculation:
  - $\mathcal{Y}$ is degraded to $\mathcal{X}$ to form pair
  - FR model estimate $\hat{\mathcal{Y}}$ based on $\mathcal{X}$
  - Calculate losses based on $\hat{\mathcal{Y}}$, $\mathcal{Y}$ and attribute
- Model only trained with paired image and attribute label

Pseudo-pair Strategy



- ➢ $\hat{\mathcal{Y}}$ is degraded to $\hat{\mathcal{X}}$ by the same degradation model
- ➢ Calculate losses based on $\hat{\mathcal{X}}$, $\mathcal{X}$ and attribute ($L_{deg} = |\mathcal{X} - \hat{\mathcal{X}}|$)
- ➢ Straight-through estimator for gradient calculation

Metrics

- Information fidelity:
  - PSNR, SSIM

$$\mathrm{SSIM}(x,y)=\frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)}$$

- Image quality:
  - NIQE, FID

$$PSNR=10\cdot\log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$

- Attribute Bias:

$$\text{Age/Gender error}=1-\frac{\sum_{y\in\mathcal{D}}\mathbb{I}(\mathbf{C}(\hat{y})\neq\mathbf{C}(y))}{|\mathcal{D}|}$$

$\mathbf{C}$: Pretrained Classifier
c_1 c_2 : Constants
MAX_I^2 : 255^2 or 1^2

Performance Comparison

➢Quantitatively comparison on CelebA-HQ (Synthesized LQs):

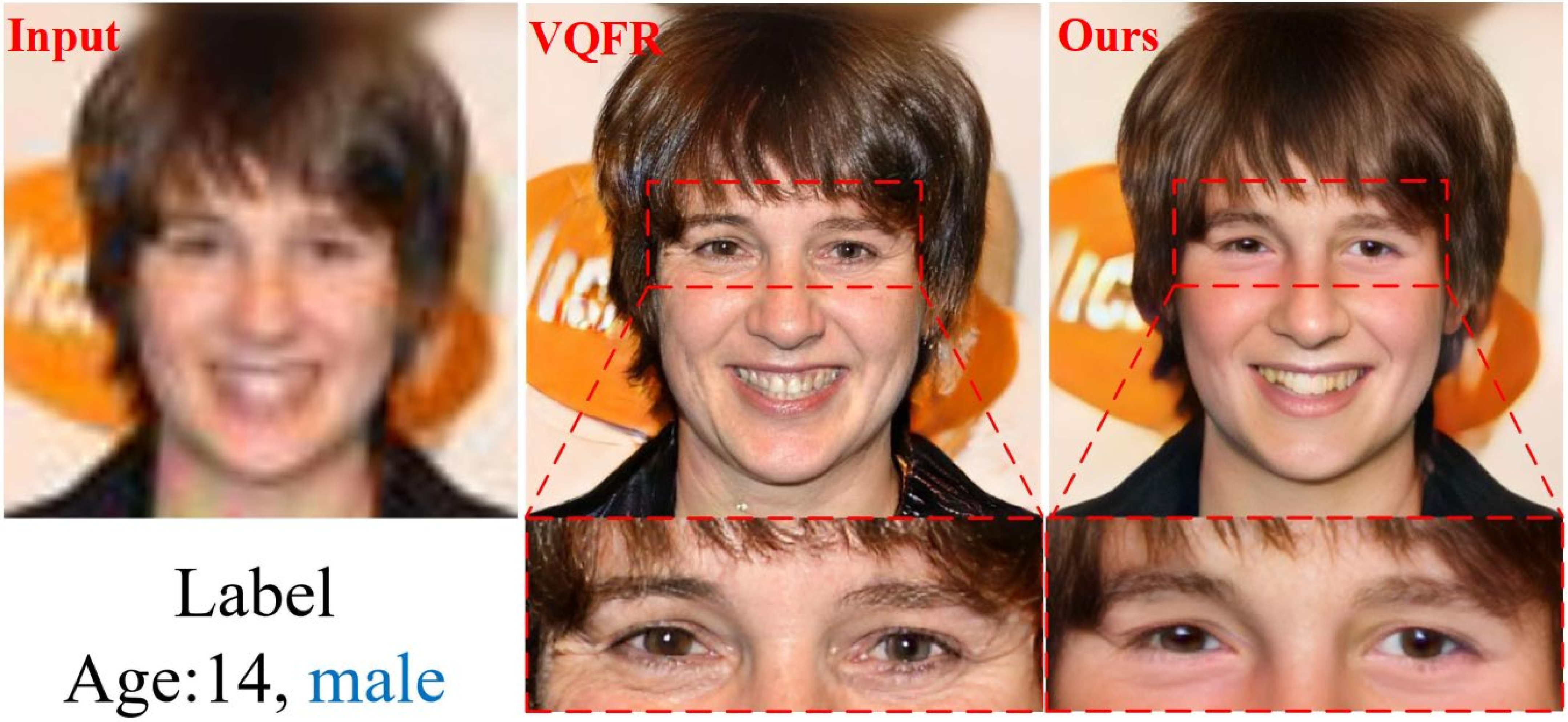
CelebA-HQ	NIQE↓	FID↓	PSNR↑	SSIM↑	Age error (%)↓	Gender error (%)↓
Bilinear	16.59	213.3	**23.25**	**0.6951**	60.90	14.37
AACNN	4.132	59.80	22.65	0.6052	43.60	2.930
PULSE	3.765	65.90	20.81	0.5695	68.90	14.16
PSFRGAN	3.982	55.88	21.71	0.6173	50.66	3.500
GFPGAN	3.800	53.87	20.72	0.6001	50.50	4.160
GPEN	3.877	**47.39**	22.12	0.6152	47.70	3.333
VQFR	**3.350**	52.09	20.48	0.5699	50.60	3.666
Ours	4.411	48.63	21.71	0.6247	**22.30**	**1.467**

The best, second results are in **red** and blue respectively.

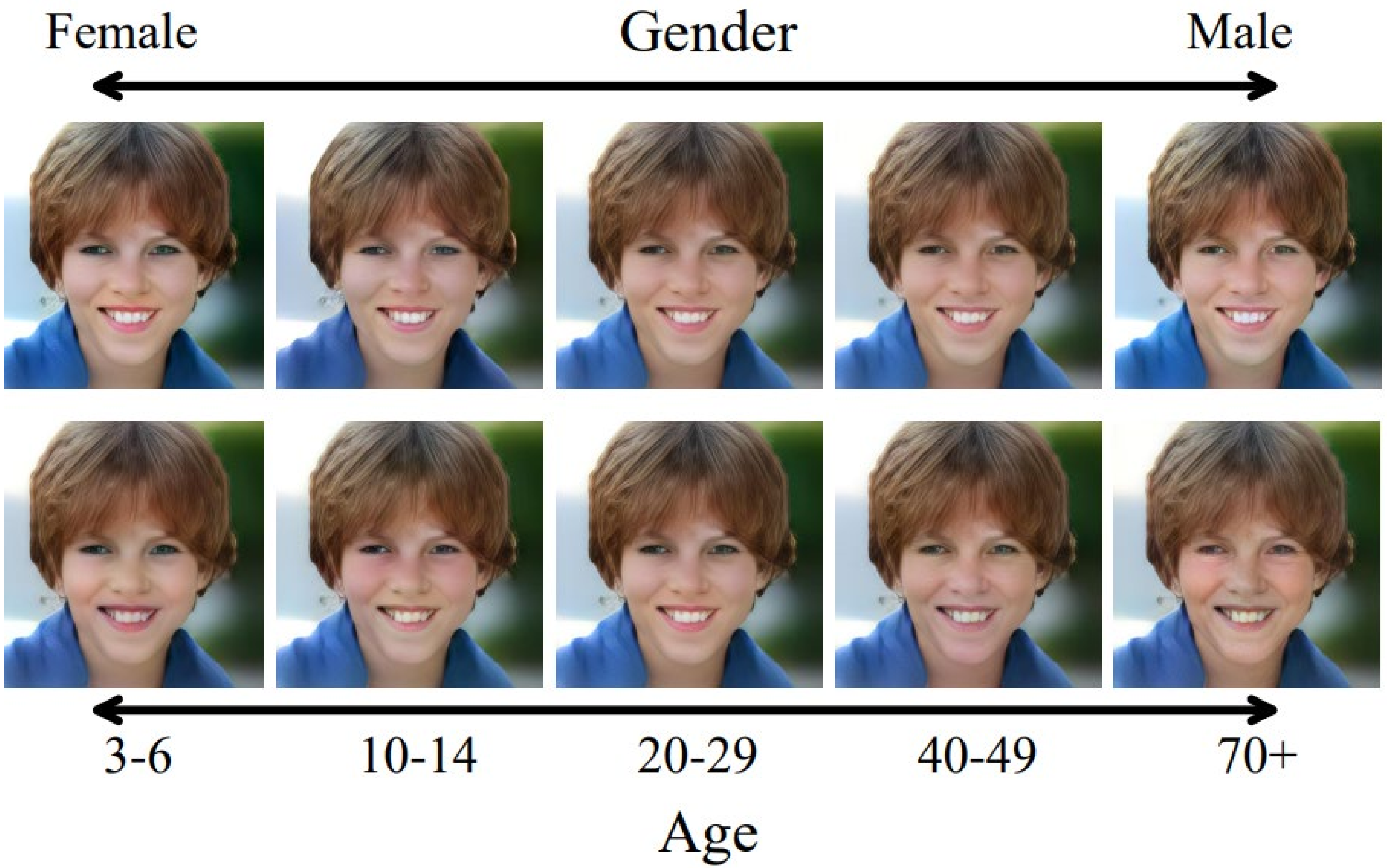
➤Quantitatively comparison on IMDB&COX (Captured LQs):

IMDB	NIQE↓	FID↓	Age error (%)↓	Gender error (%)↓
PSFRGAN	4.316	39.04	50.72	1.790
GFPGAN	4.133	**32.30**	48.30	1.699
GPEN	4.719	54.13	49.25	1.155
VQFR	**3.540**	33.40	50.29	1.631
Ours	4.482	39.41	**26.57**	**0.929**
COX				
PSFRGAN	4.521	88.94	61.60	21.19
GFPGAN	5.036	82.52	59.50	20.59
GPEN	4.713	84.01	64.30	22.59
VQFR	**4.190**	**70.00**	59.40	22.90
Ours	5.238	80.45	**29.50**	**8.80**

The best, second results are in **red** and blue respectively.


Input
VQFR
Ours
Label
Age:14, male

Human Interaction



Summarization

- **Attribute bias problem**: facial attributes (e.g., age and gender) of the restored faces could be dramatically different from the target faces

- Owe it to two leading causes: *the lack of attribute information* and *bias in training data*

- Propose **DebiasFR**, which faithfully preserves input attribute information and produces quality HQ faces. Supports human interaction attribute adjustment

Thanks

Presenter: Zelin Li

12132338@mail.sustech.edu.cn