

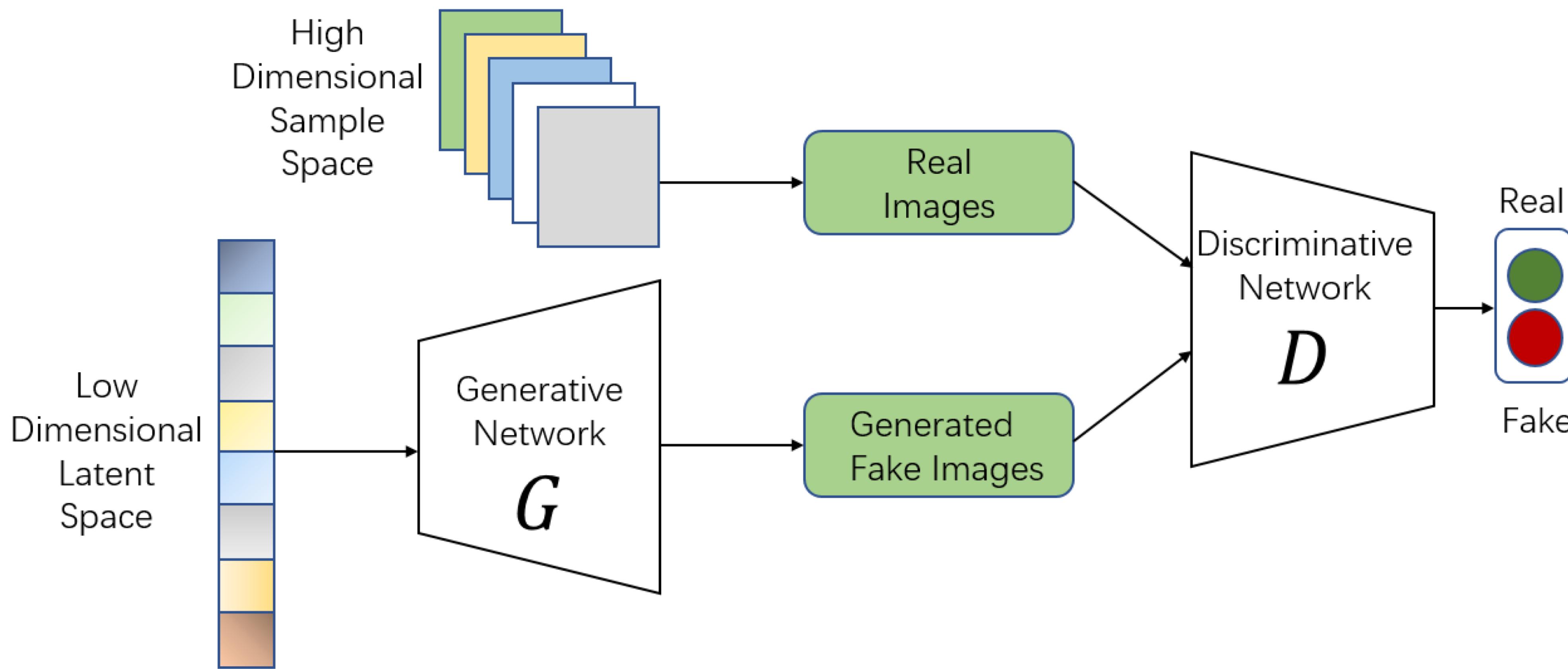


НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

# ANYCOST GAN

Василевская Юля  
БПМИ191

$$\min_G \max_D V(D, G) = \underset{x \sim p_{data}}{E} [\log D(x)] + \underset{z \sim p_z}{E} [\log(1 - D(G(z)))]$$



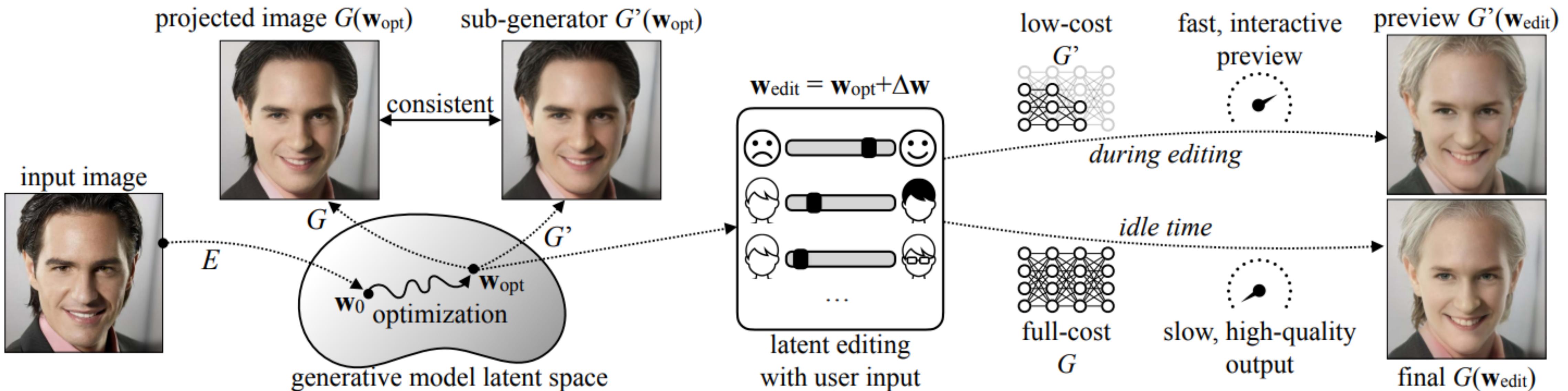
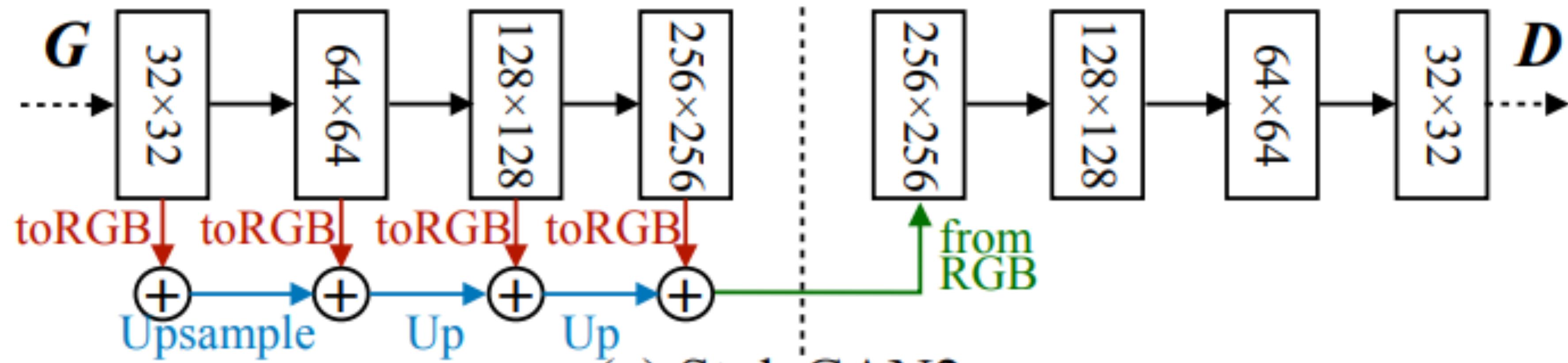
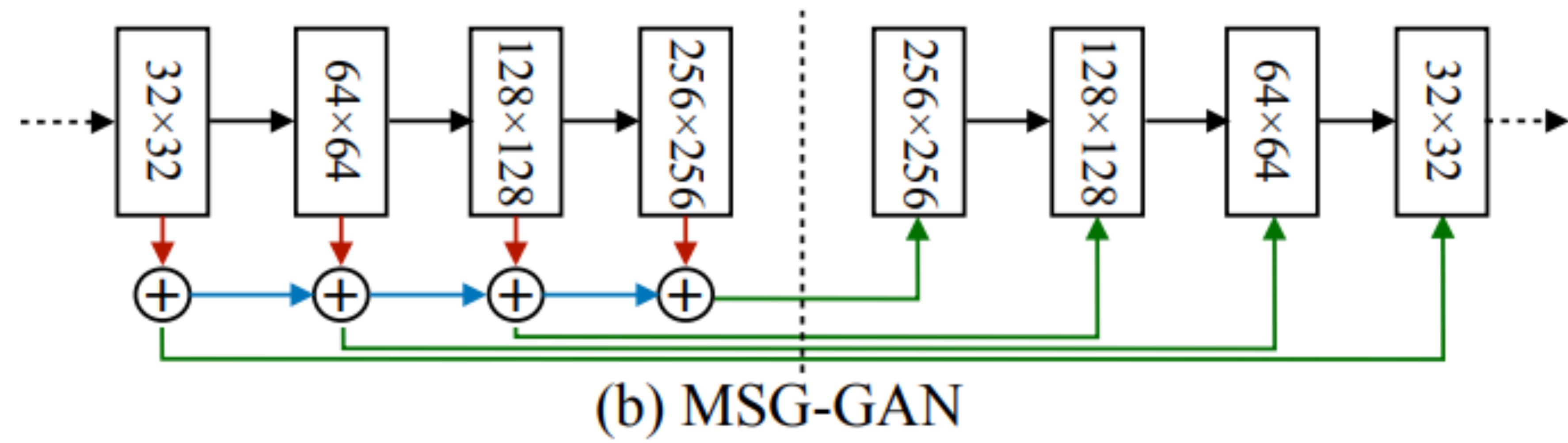


Figure 2: Anycost GAN for image synthesis and editing. Given an input image, we project it into the latent space with encoder  $E$  and backward optimization. We can modify the latent code with user input to edit the image. During editing, a sub-generator of small cost is used for fast and interactive preview; during idle time, the full cost generator renders the final, high-quality output. The outputs from the full and sub-generators are visually consistent during projection and editing.



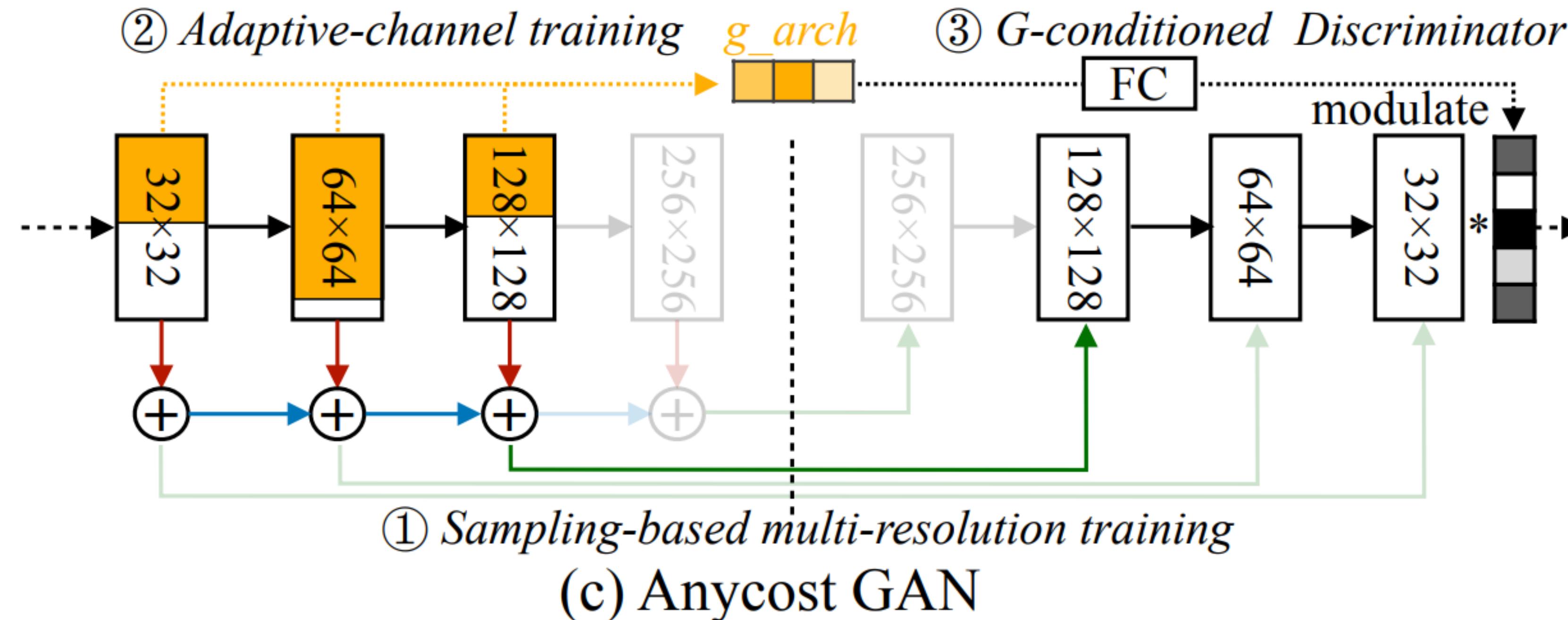
(a) StyleGAN2



(b) MSG-GAN

# Multi-resolution training

# Anycost GANs



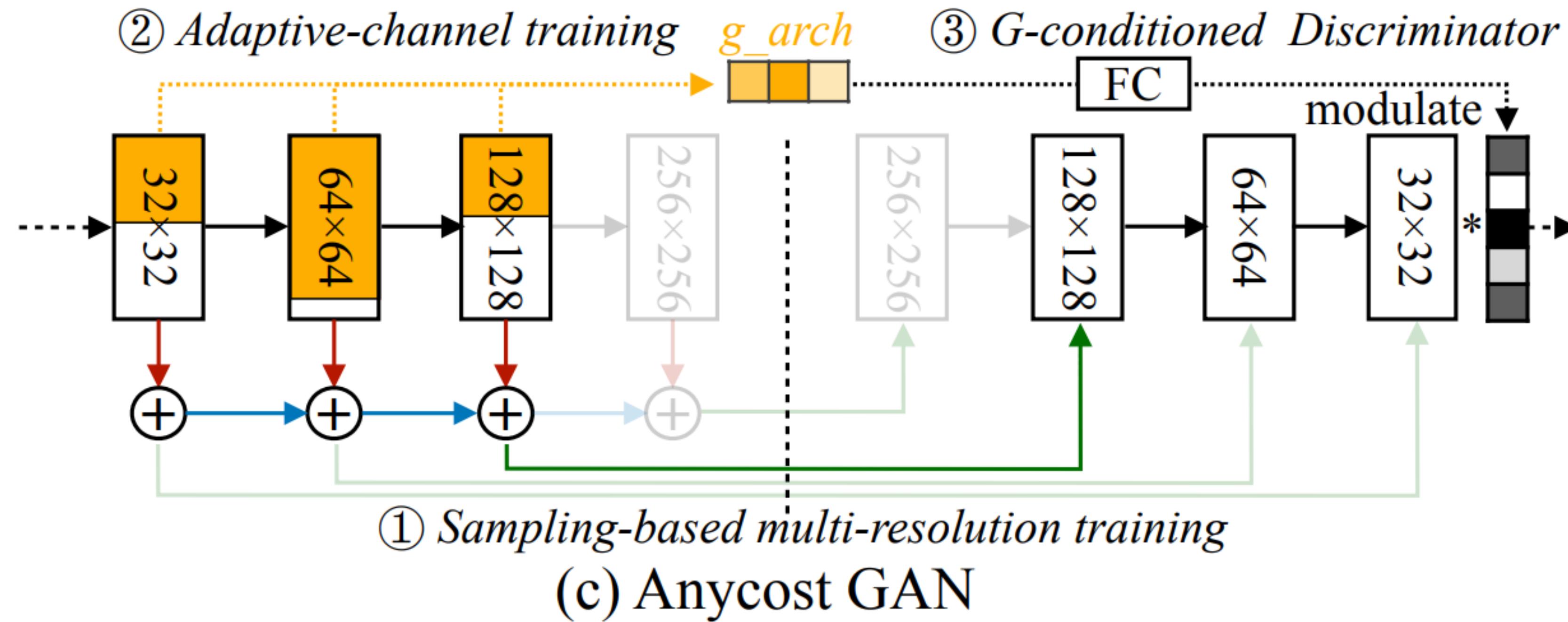
$$\tilde{\mathbf{x}} = G(\mathbf{w}) = g^K \circ g^{K-1} \circ \dots \circ g^k \circ \dots \circ g^2 \circ g^1(\mathbf{w})$$

$$\tilde{\mathbf{x}}^k = G^k(\mathbf{w}) = g^k \circ g^{k-1} \circ \dots \circ g^2 \circ g^1(\mathbf{w}), k \leq K$$

$$\mathcal{L}_{\text{multi-res}} = \mathbb{E}_{\mathbf{x}, k} [\log D(\mathbf{x}^k)] + \mathbb{E}_{\mathbf{w}, k} [\log(1 - D(G^k(\mathbf{w})))]$$

# Adaptive-channel training

# Anycost GANs

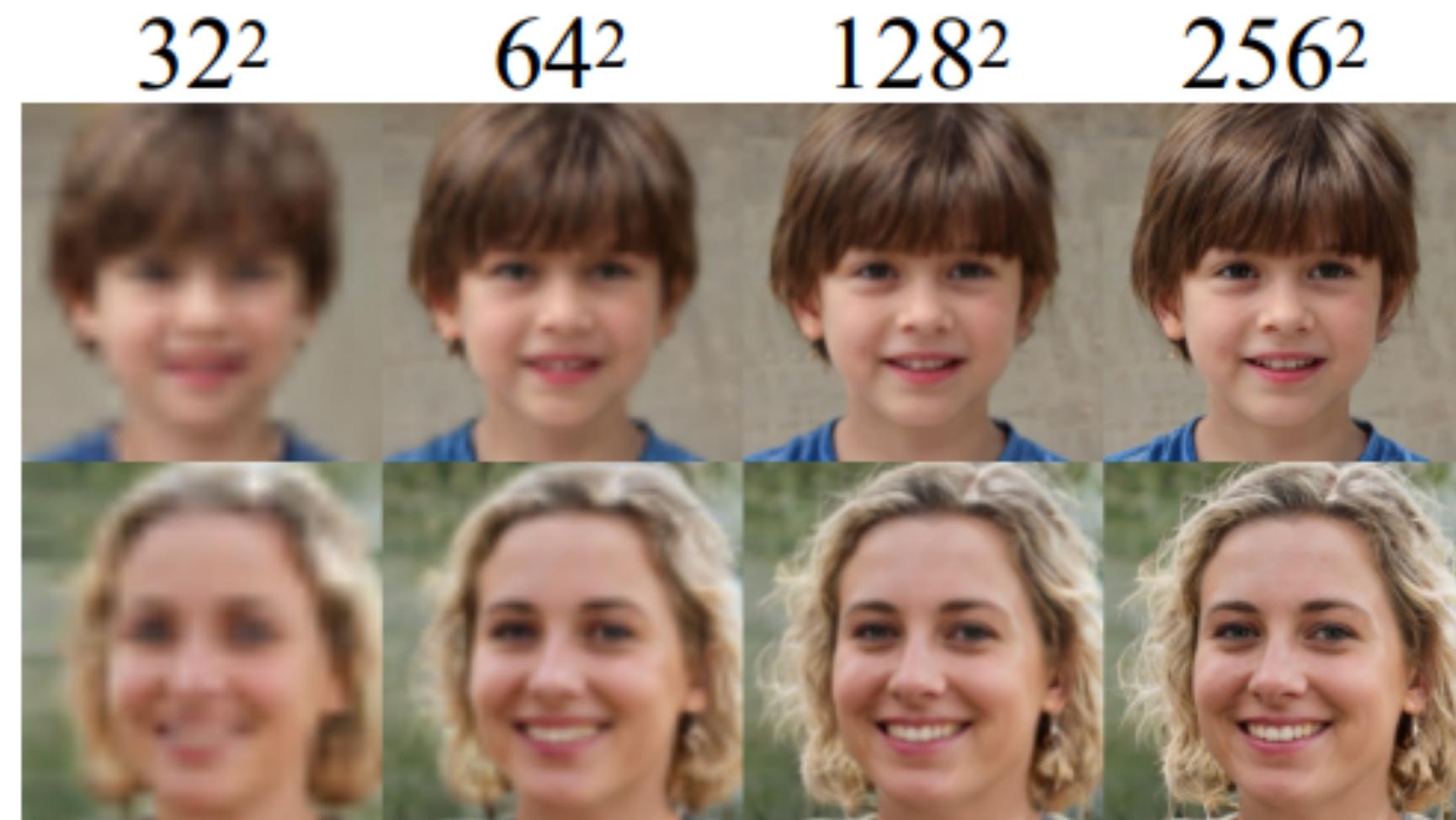
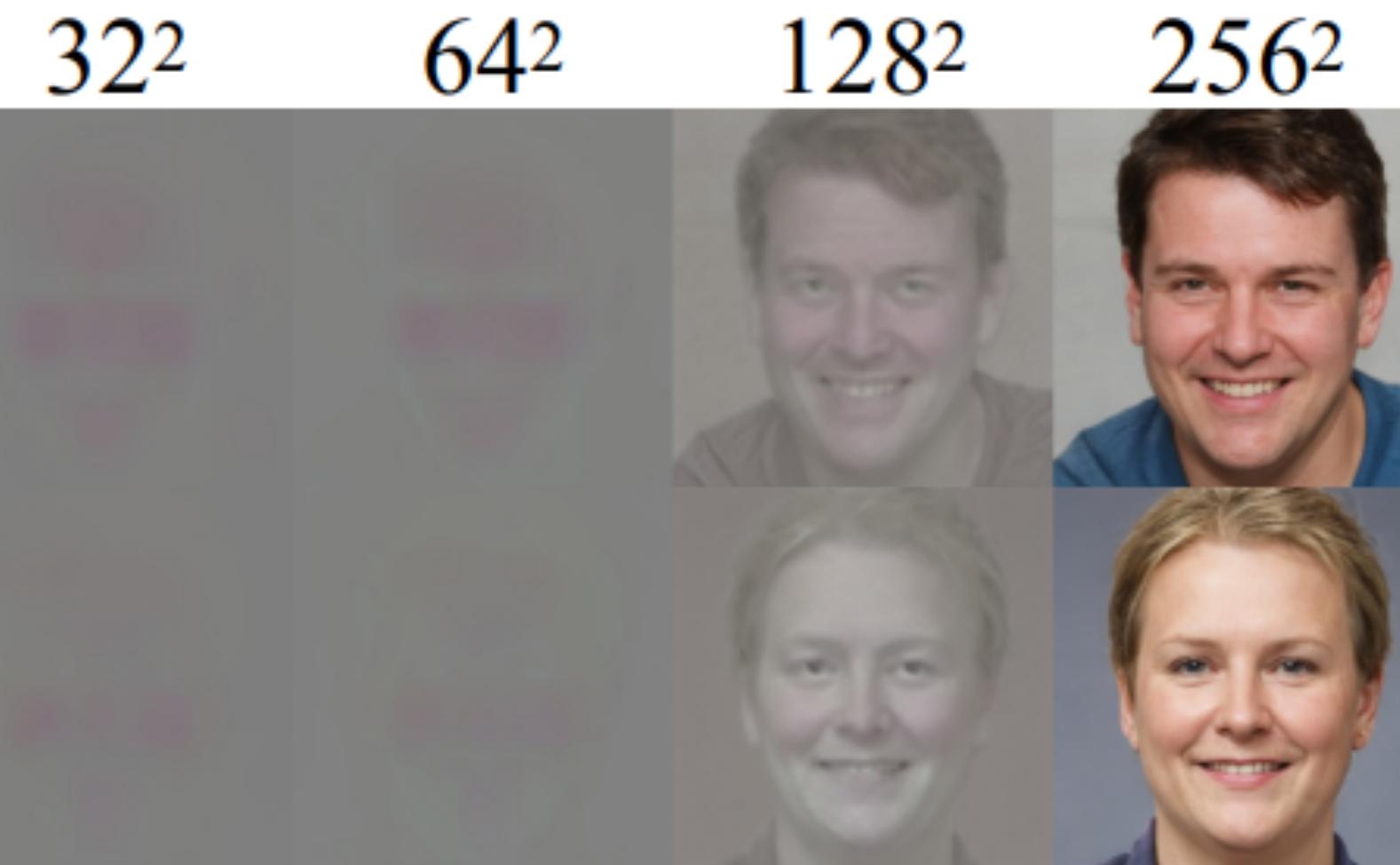


$$\mathcal{L}_{\text{ada-ch}} = \mathbb{E}_{\mathbf{x}, k} [\log D(\mathbf{x}^k)] + \mathbb{E}_{\mathbf{w}, k, \mathbb{C}} [\log(1 - D(G_{\mathbb{C}}^k(\mathbf{w})))]$$

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{ada-ch}} + \mathbb{E}_{\mathbf{w}, k, \mathbb{C}} [\ell(G_{\mathbb{C}}^k(\mathbf{w}), G(\mathbf{w}))]$$

# Results

## Anycost GANs

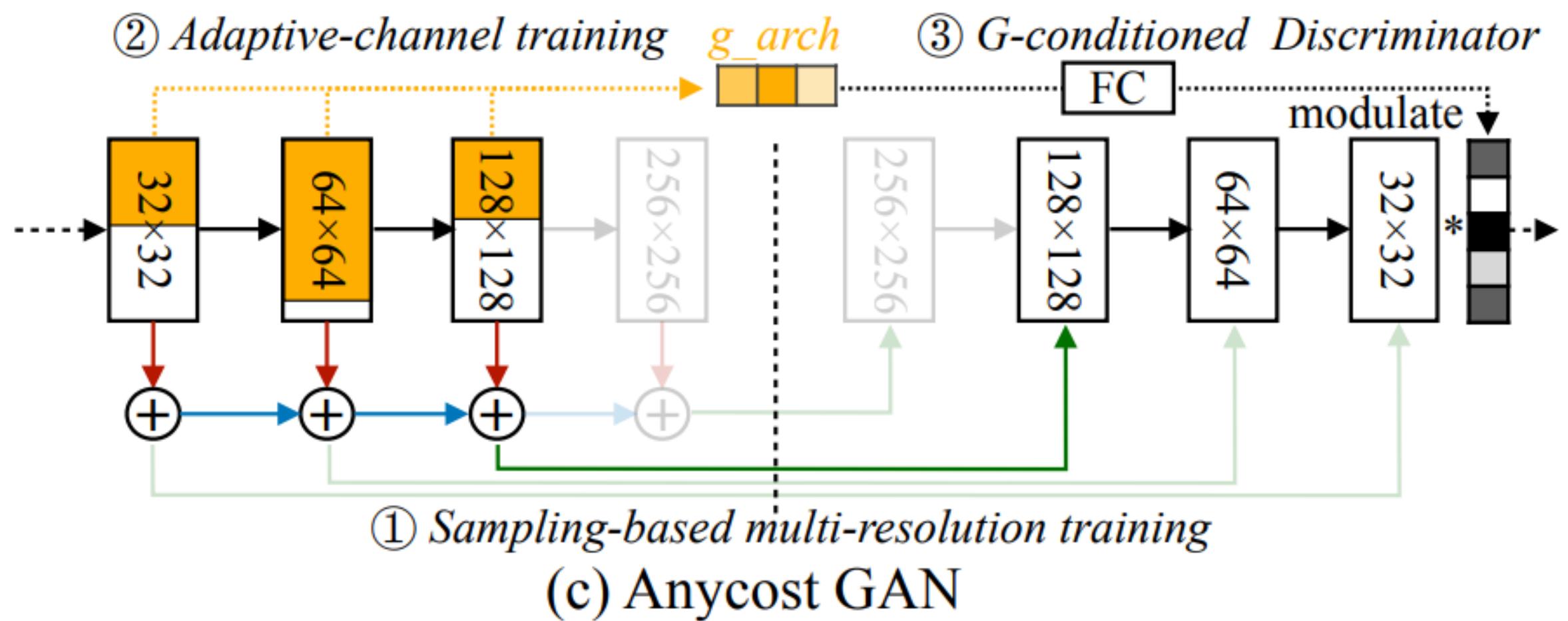


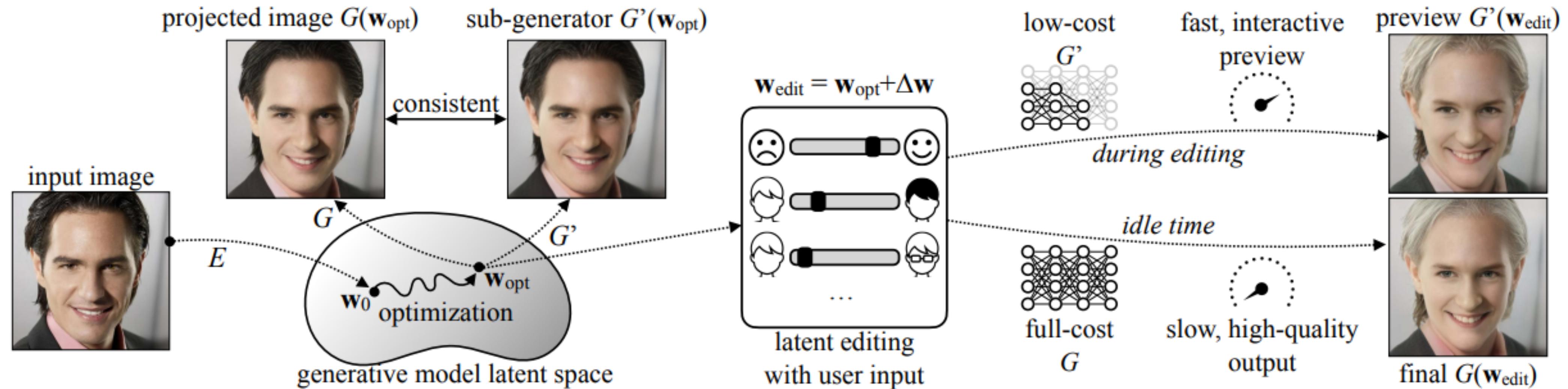
# Generator-conditioned discriminator

# Anycost GANs

Table 2: FIDs on FFHQ at different resolutions and channels. All the settings except “vanilla” only train one generator and evaluate it at multiple configurations. We mark a better FID **green** and a worse FID **red** compared to multi-G baseline. Conditioned discriminator (“conditioned”) provides the best FID over different channel widths and resolutions. The model is based on Config-E for faster ablation.

FID-70k↓	resolution 256				resolution 128			
	1×	0.75×	0.5×	0.25×	1.0×	0.75×	0.5×	0.25×
vanilla	3.80	4.64	6.20	10.39	4.04	4.99	5.78	11.15
same D	3.63	3.91	5.41	14.01	7.25	6.81	5.92	7.57
reduced ch	3.67	4.35	5.82	10.62	3.09	3.65	4.74	8.82
conditioned (Ours)	3.73	3.86	4.64	8.06	3.30	3.28	3.69	5.55





$$E^* = \arg \min_E \mathbb{E}_{\mathbf{x}} [\ell(G(E(\mathbf{x})), \mathbf{x}) + \alpha \mathbb{E}_{k, \mathcal{C}} \ell(G_{\mathcal{C}}^k(E(\mathbf{x})), \mathbf{x})]$$

$$\mathbf{w}^* = \arg \min_{\mathbf{w}} [\ell(G(\mathbf{w}), \mathbf{x}) + \alpha \mathbb{E}_{k, \mathcal{C}} \ell(G_{\mathcal{C}}^k(\mathbf{w}, \mathbf{x})]$$

# Comparison

# Anycost GANs

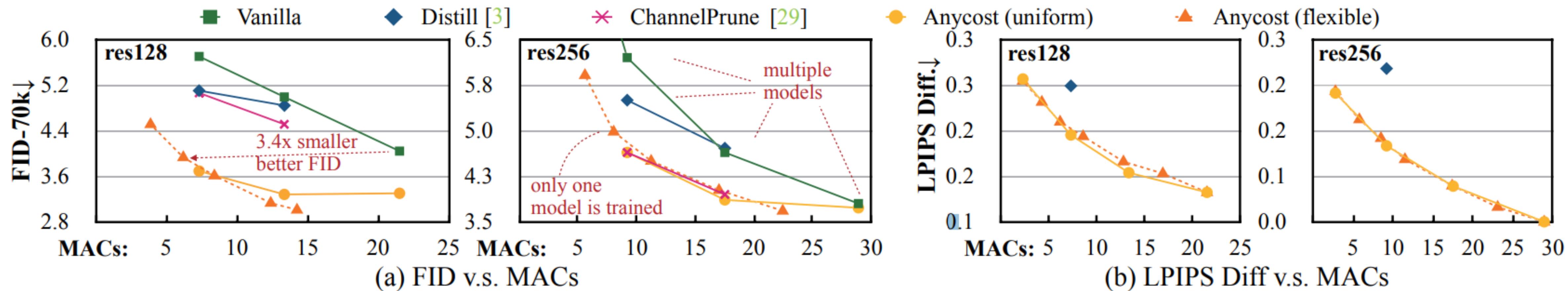


Figure 5: Anycost GAN outperforms existing compression baselines [3, 29] at various computation budgets, despite only training a single, flexible generator across computation budgets.

Table 3: Anycost GANs achieves similar or better FIDs/path lengths at various channel widths compared to StyleGAN2, despite training a single, flexible generator.

Channels:	FID-50k↓				Path length↓	
	1.0×	0.75×	0.5×	0.25×	1.0×	0.5×
FFHQ	StyleGAN2	2.84	-	3.31	-	145.0
	Anycost	2.77	3.05	3.28	5.01	144.2
1024	StyleGAN2	2.32	-	3.19	-	415.5
	Anycost	2.38	2.46	2.61	3.69	471.2
Car	StyleGAN2	2.32	-	3.19	-	415.5
	Anycost	2.38	2.46	2.61	3.69	430.0

	Smiling	BlackHair	Eyeglass	StraightHair	Earrings
Separate	55.8%	62.7%	83.4%	55.8%	50.6%
0.5× ch	98.1%	97.0%	99.9%	95.9%	94.1%
0.25× ch	96.9%	95.5%	99.8%	93.6%	88.4%

Table 1: FID-70k on FFHQ of different multi-resolution training techniques. Our sampling-based technique can train one model that produces multiple resolution outputs with higher image quality (measured by FID [30]) compared to single resolution training. The models are trained with half channels (Config-E) for faster ablation.

Resolution	1024	512	256	128	64	32
Single-resolution	3.25	4.17	3.76	4.04	3.32	2.41
MSG-GAN [41]	-	-	4.79	6.34	2.7	3.04
Ours (low res)	-	-	3.49	<b>3.26</b>	<b>2.52</b>	<b>2.18</b>
Ours (high res)	<b>2.99</b>	<b>3.08</b>	<b>3.35</b>	3.98	-	-

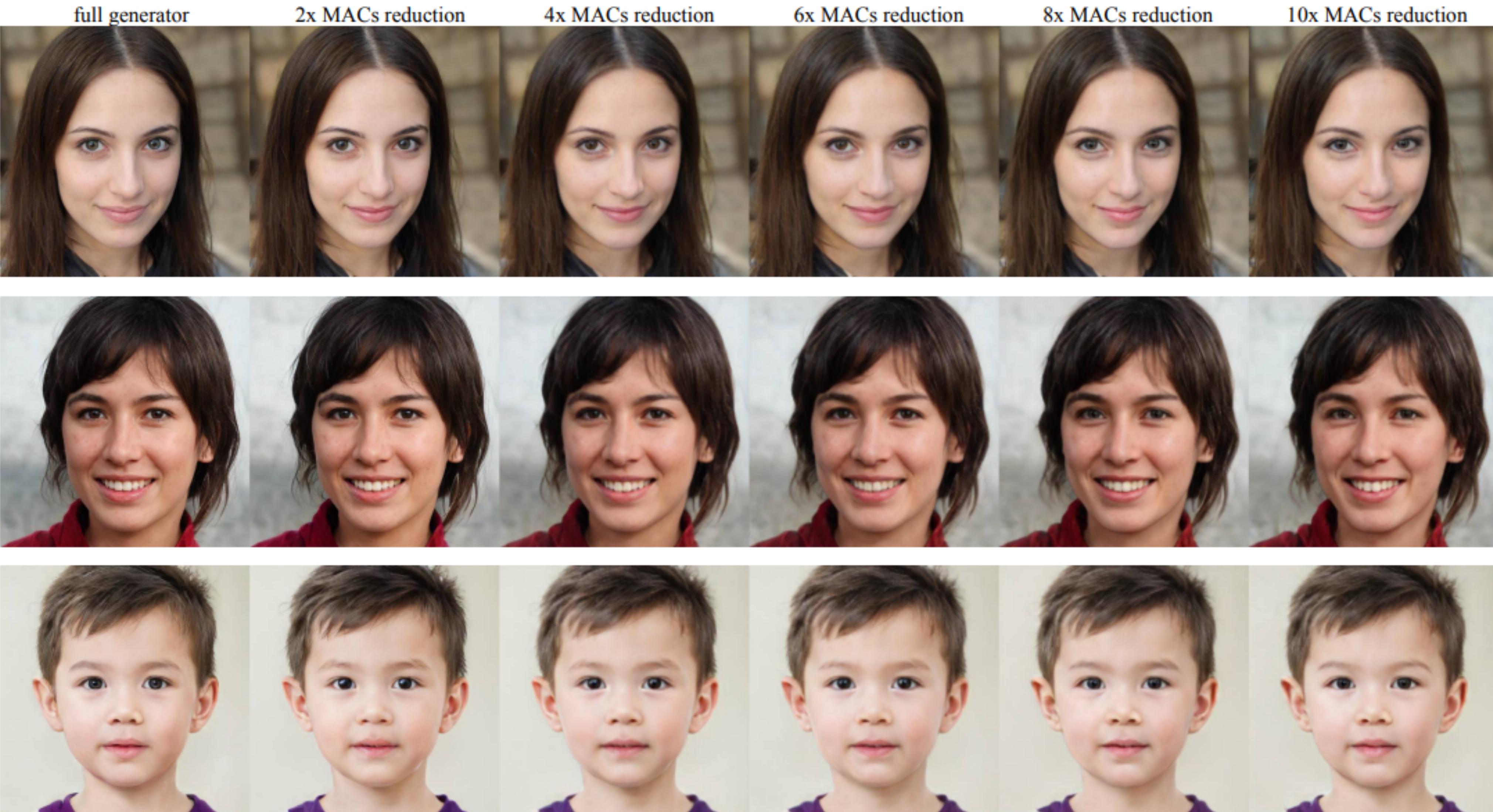
# Uniform on FFHQ

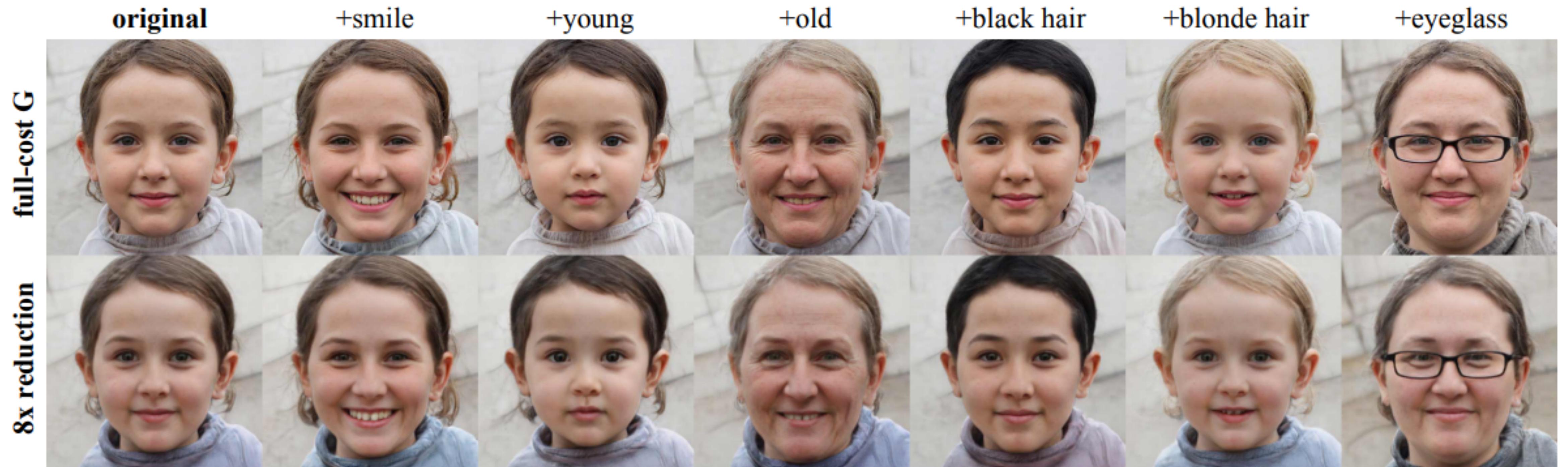
# Anycost GANs



# Flexible on FFHQ

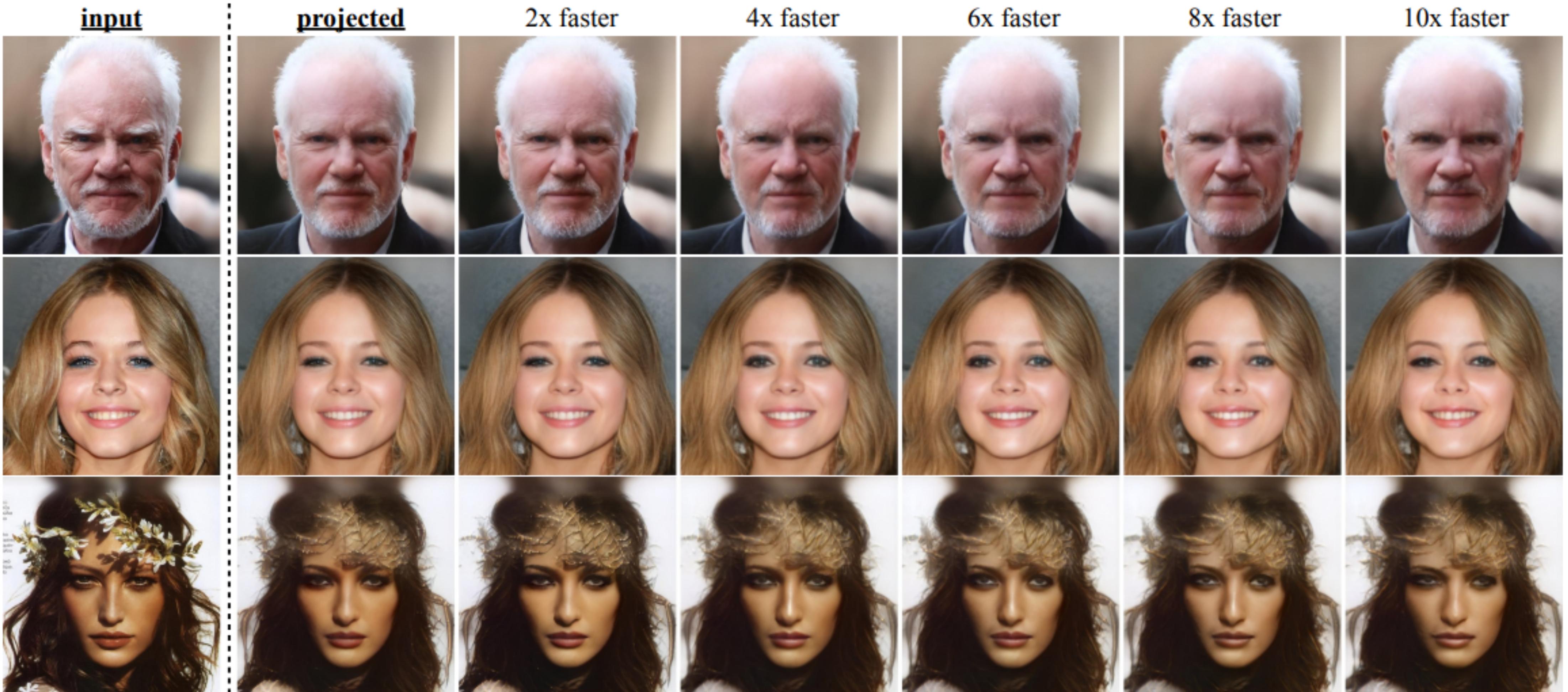
# Anycost GANs





# Projection for FFHQ

Anycost GANs



*Кажется, на этом всё  
Надеюсь, было интересно :)*