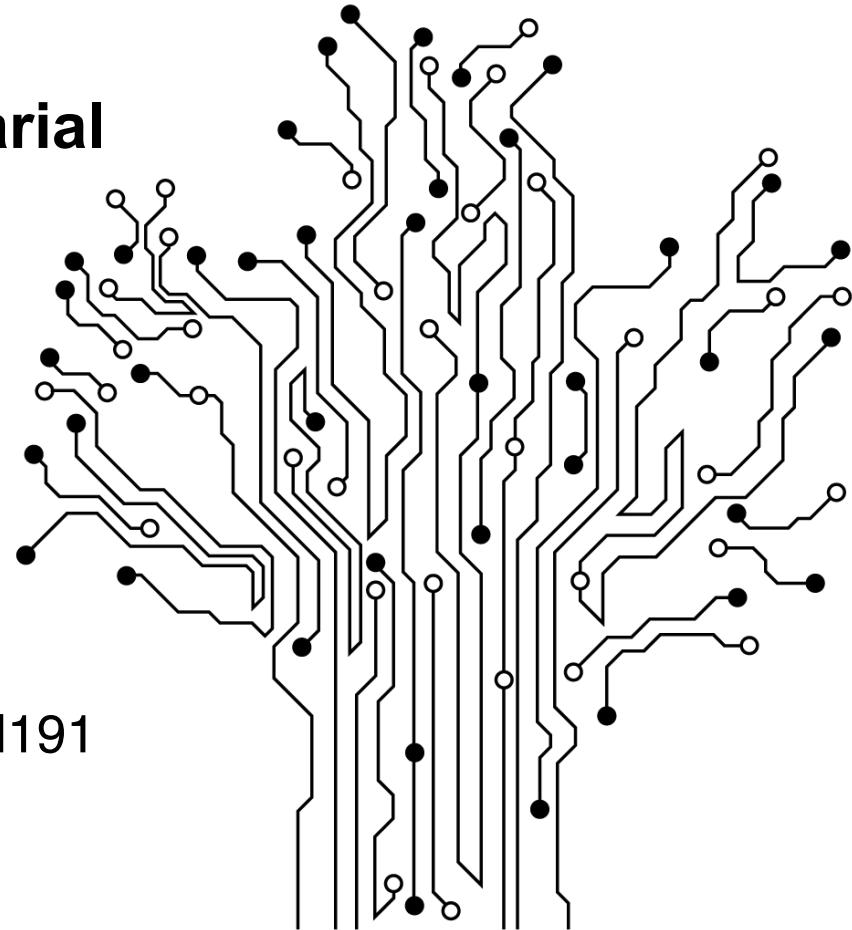


StyleGAN3

Alias-Free Generative Adversarial Networks

Подготовила Курченко Лилия, БПМИ191



StyleGAN2



- Универсален – работает с любыми подходящими данными
- Гиперреалистичное качество и высокое разрешение (1024*1024)
- Не подходит для видео или анимации

Проблемы StyleGAN2



- Изображения на самом деле генерируются неиерархически
- Прилипание текстуры (детали привязываются к их координатам в пикселях, унаследованным от обучения)

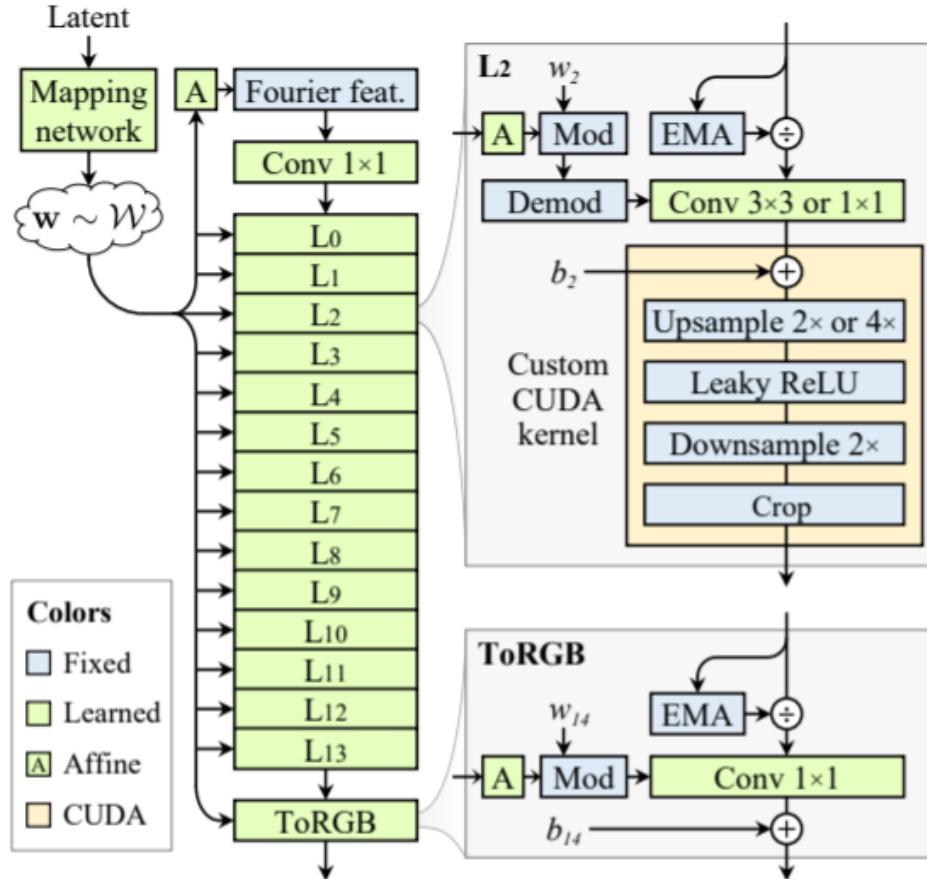
<https://youtu.be/-2hLdOonvK0>

Что предлагает StyleGAN3



- Изменение обработки сигналов
- переход к интерпретации всех сигналов в нейронной сети как непрерывных процессов
- генератор становится эквивариантным сдвигу и вращению

Архитектура



Архитектура

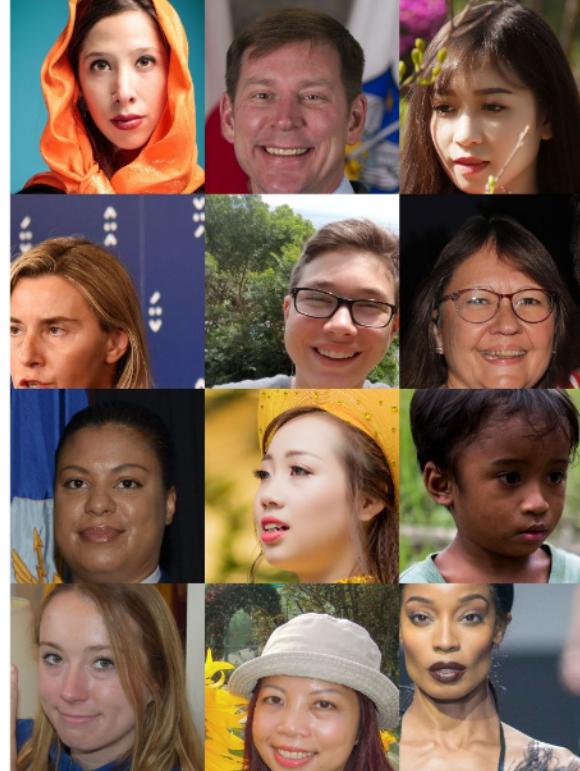


<https://youtu.be/Lu40bcm4YtY>

FFHQ и FFHQ -U



FFHQ



FFHQU

Результаты для FFHQ-U



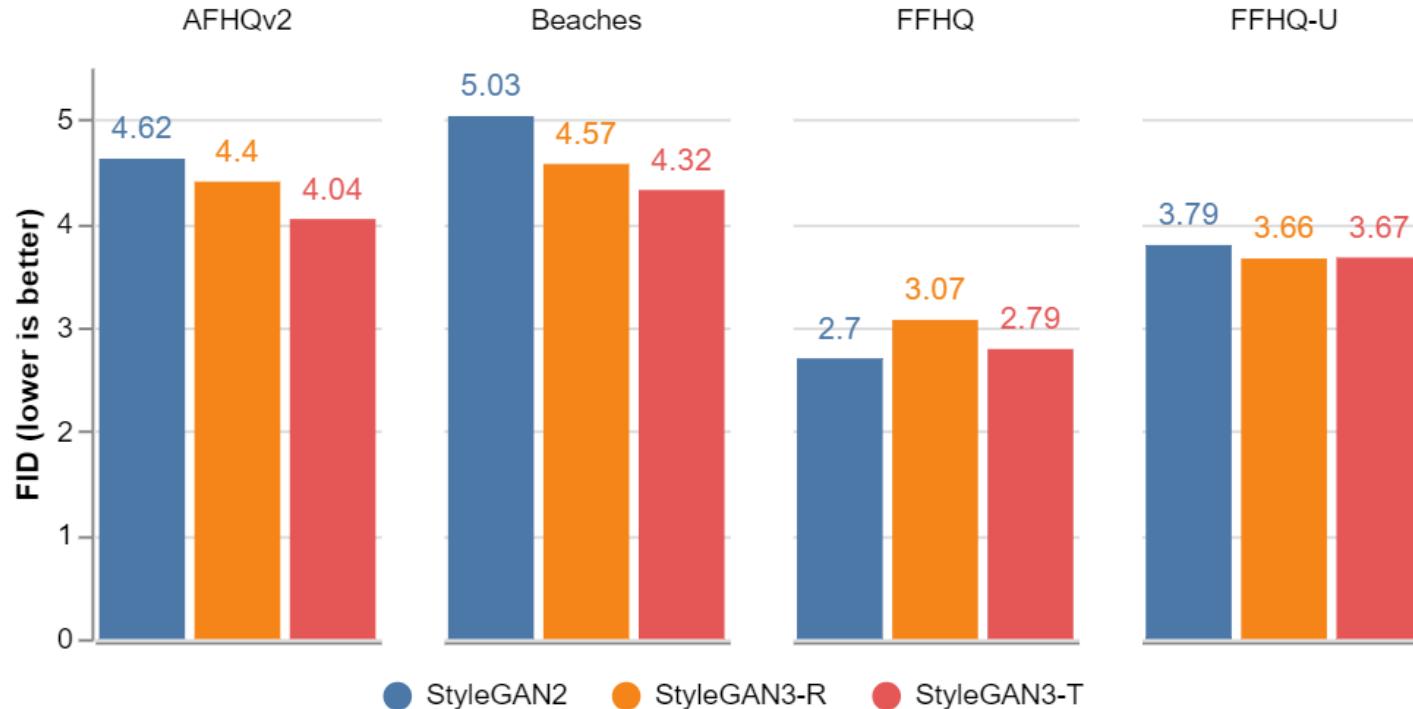
Configuration	FID ↓	EQ-T ↑	EQ-R ↑
A StyleGAN2	5.14	–	–
B + Fourier features	4.79	16.23	10.81
C + No noise inputs	4.54	15.81	10.84
D + Simplified generator	5.21	19.47	10.41
E + Boundaries & upsampling	6.02	24.62	10.97
F + Filtered nonlinearities	6.35	30.60	10.81
G + Non-critical sampling	4.78	43.90	10.84
H + Transformed Fourier features	4.64	45.20	10.61
T + Flexible layers (StyleGAN3-T)	4.62	63.01	13.12
R + Rotation equiv. (StyleGAN3-R)	4.50	66.65	40.48

Обучающие конфигурации.

FID вычисляется между 50 тыс. сгенерированных изображений и всеми обучающими изображениями; чем меньше, тем лучше.

EQ-T и EQR - это наши показатели эквивариантности в дБ; чем выше, тем лучше.

Результаты



Результаты

Dataset	Config	FID ↓	EQ-T ↑	EQ-R ↑
FFHQ-U 70000 img, 1024 ² Train from scratch	StyleGAN2	3.79	15.89	10.79
	StyleGAN3-T (ours)	3.67	61.69	13.95
	StyleGAN3-R (ours)	3.66	64.78	47.64
FFHQ 70000 img, 1024 ² Train from scratch	StyleGAN2	2.70	13.58	10.22
	StyleGAN3-T (ours)	2.79	61.21	13.82
	StyleGAN3-R (ours)	3.07	64.76	46.62
METFACES-U 1336 img, 1024 ² ADA, from FFHQ-U	StyleGAN2	18.98	18.77	13.19
	StyleGAN3-T (ours)	18.75	64.11	16.63
	StyleGAN3-R (ours)	18.75	66.34	48.57
METFACES 1336 img, 1024 ² ADA, from FFHQ	StyleGAN2	15.22	16.39	12.89
	StyleGAN3-T (ours)	15.11	65.23	16.82
	StyleGAN3-R (ours)	15.33	64.86	46.81
AFHQv2 15803 img, 512 ² ADA, from scratch	StyleGAN2	4.62	13.83	11.50
	StyleGAN3-T (ours)	4.04	60.15	13.51
	StyleGAN3-R (ours)	4.40	64.89	40.34
BEACHES 20155 img, 512 ² ADA, from scratch	StyleGAN2	5.03	15.73	12.69
	StyleGAN3-T (ours)	4.32	59.33	15.88
	StyleGAN3-R (ours)	4.57	63.66	37.42

Ablation	Translation eq.		+ Rotation eq.		
	FID ↓	EQ-T ↑	FID ↓	EQ-T ↑	EQ-R ↑
* Main configuration	4.62	63.01	4.50	66.65	40.48
With mixing reg.	4.60	63.48	4.67	63.59	40.90
With noise inputs	4.96	24.46	5.79	26.71	26.80
Without flexible layers	4.64	45.20	4.65	44.74	22.52
Fixed Fourier features	5.93	64.57	6.48	66.20	41.77
With path length reg.	5.00	68.36	5.98	71.64	42.18
0.5× capacity	7.43	63.14	6.52	63.08	39.89
* 1.0× capacity	4.62	63.01	4.50	66.65	40.48
2.0× capacity	3.80	66.61	4.18	70.06	42.51
* Kaiser filter, $n = 6$	4.62	63.01	4.50	66.65	40.48
Lanczos filter, $a = 2$	4.69	51.93	4.44	57.70	25.25
Gaussian filter, $\sigma = 0.4$	5.91	56.89	5.73	59.53	39.43
G-CNN comparison	FID ↓	EQ-T ↑	EQ-R ↑	Params	Time
* StyleGAN3-T (ours)	4.62	63.01	13.12	23.3M	1.00×
+ p4 symmetry [16]	4.69	61.90	17.07	21.8M	2.48×
* StyleGAN3-R (ours)	4.50	66.65	40.48	15.8M	1.37×

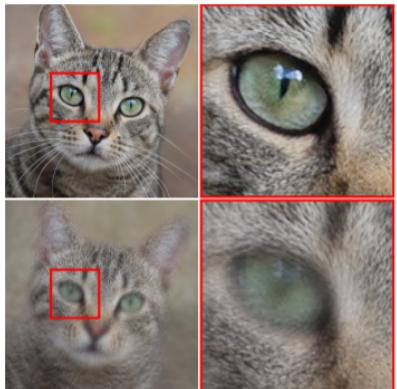
Figure 5: **Left:** Results for six datasets. We use adaptive discriminator augmentation (ADA) [32] for the smaller datasets. “StyleGAN2” corresponds to our baseline config B with Fourier features. **Right:** Ablations and comparisons for FFHQ-U (unaligned FFHQ) at 256². * indicates our default choices.

Сравнение

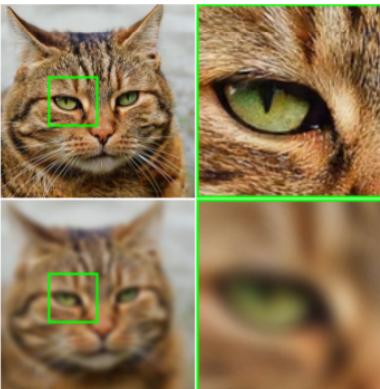


Central

StyleGAN2



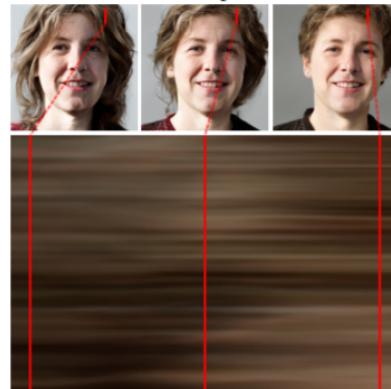
Ours



Averaged

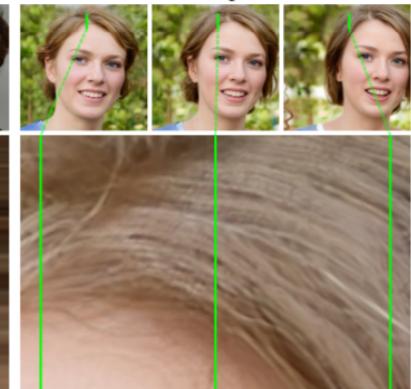
StyleGAN2

← latent interpolation →



Ours

← latent interpolation →



Сравнение



Random latent walk using directions from StyleCLIP, GANSpace, and SeFa.

Сравнение



StyleGAN2



StyleGAN3 (Ours)