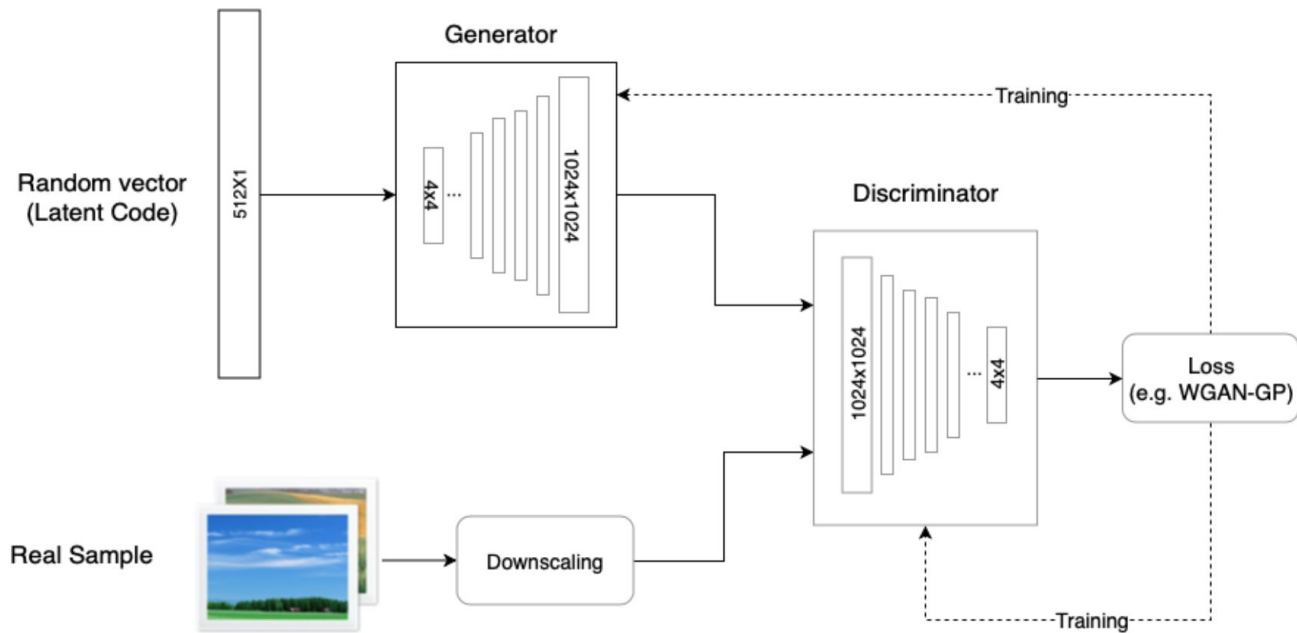


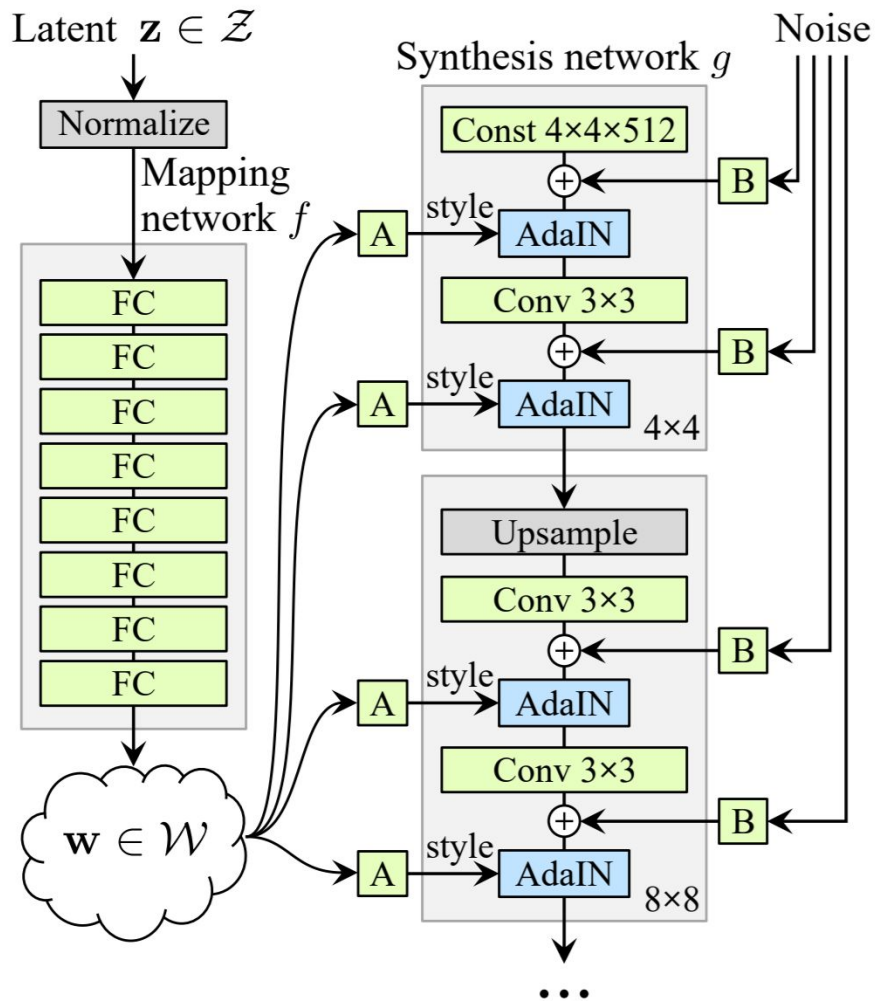
# StyleGan

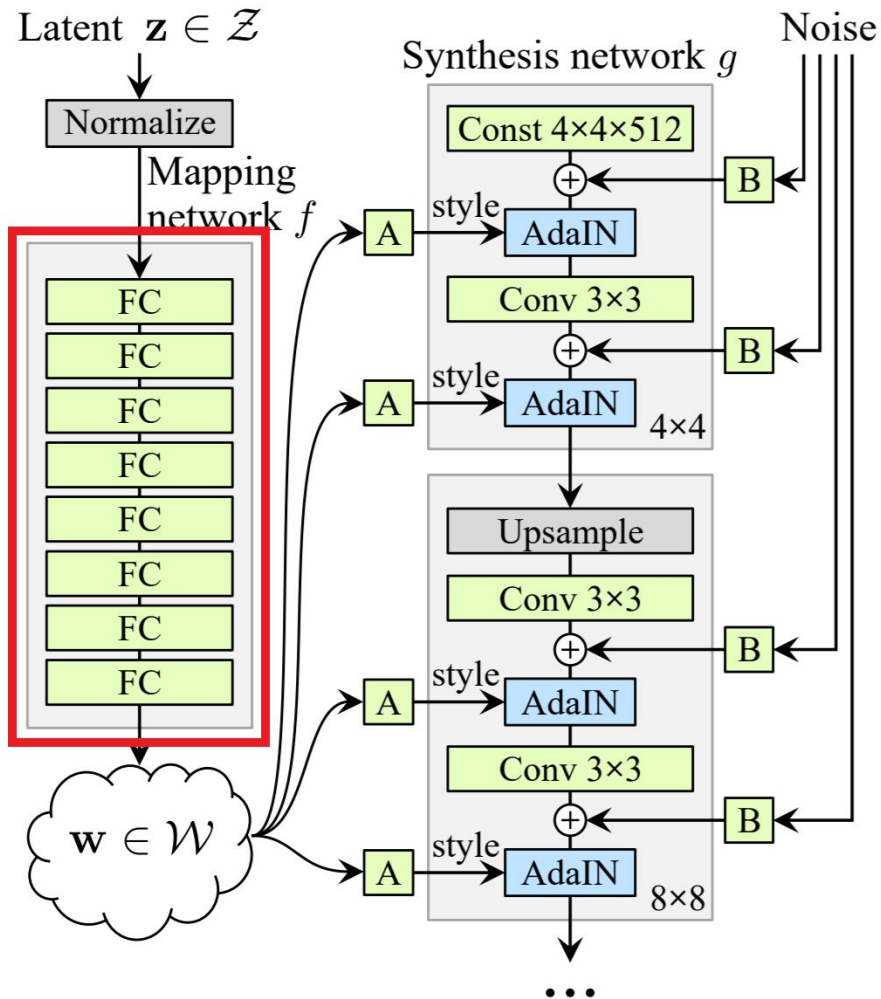
Станкевич Матвей

# ProgressiveGAN

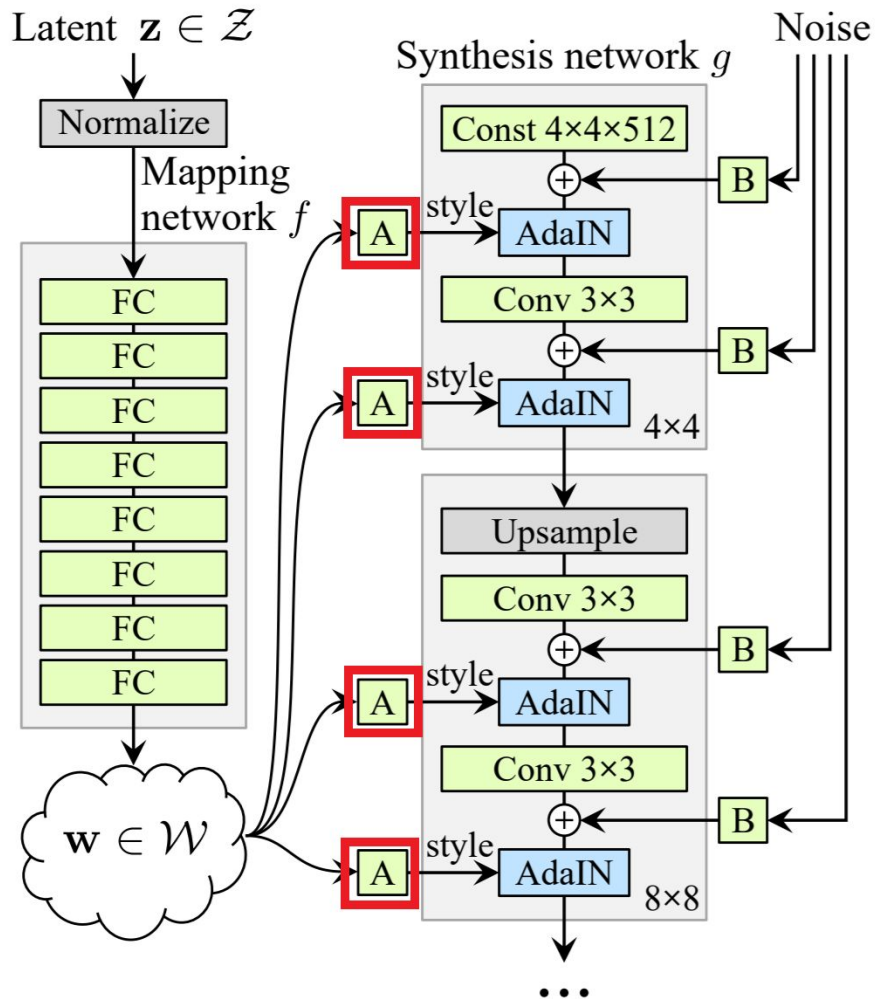


ProGAN overview



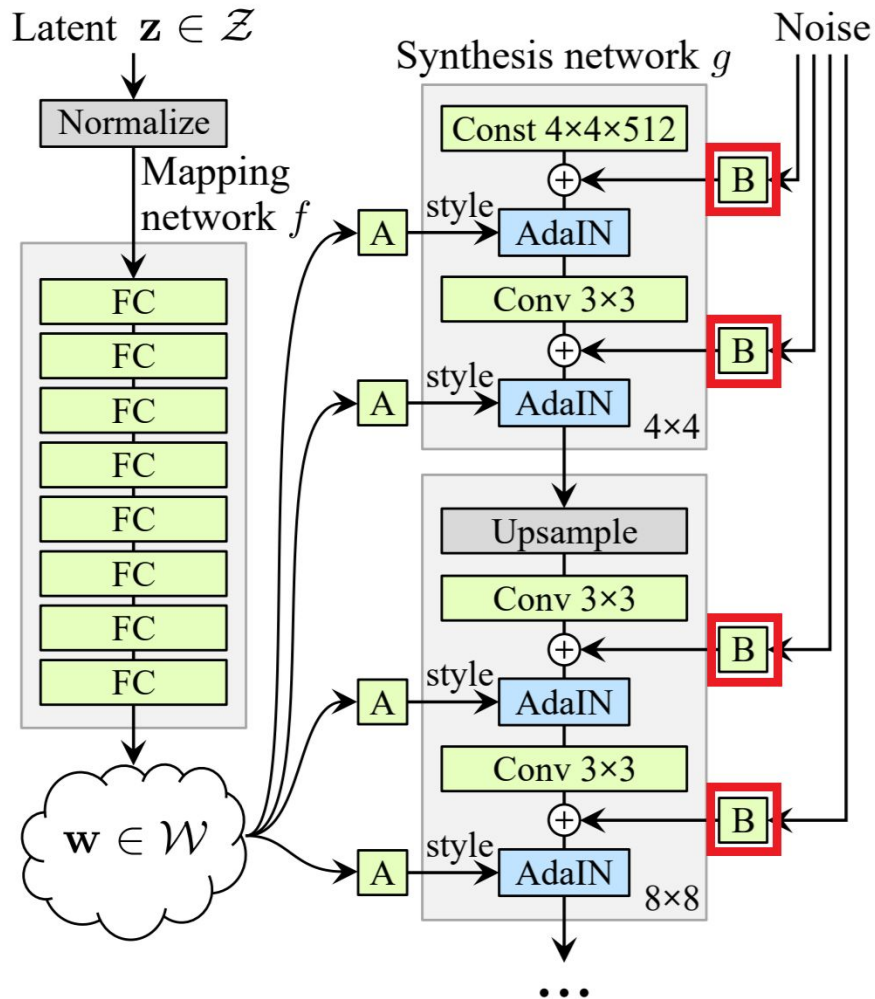


-Mapping network:  
наличие полносвязных слоёв  
помогает распутать пространство  $\mathcal{Z}$

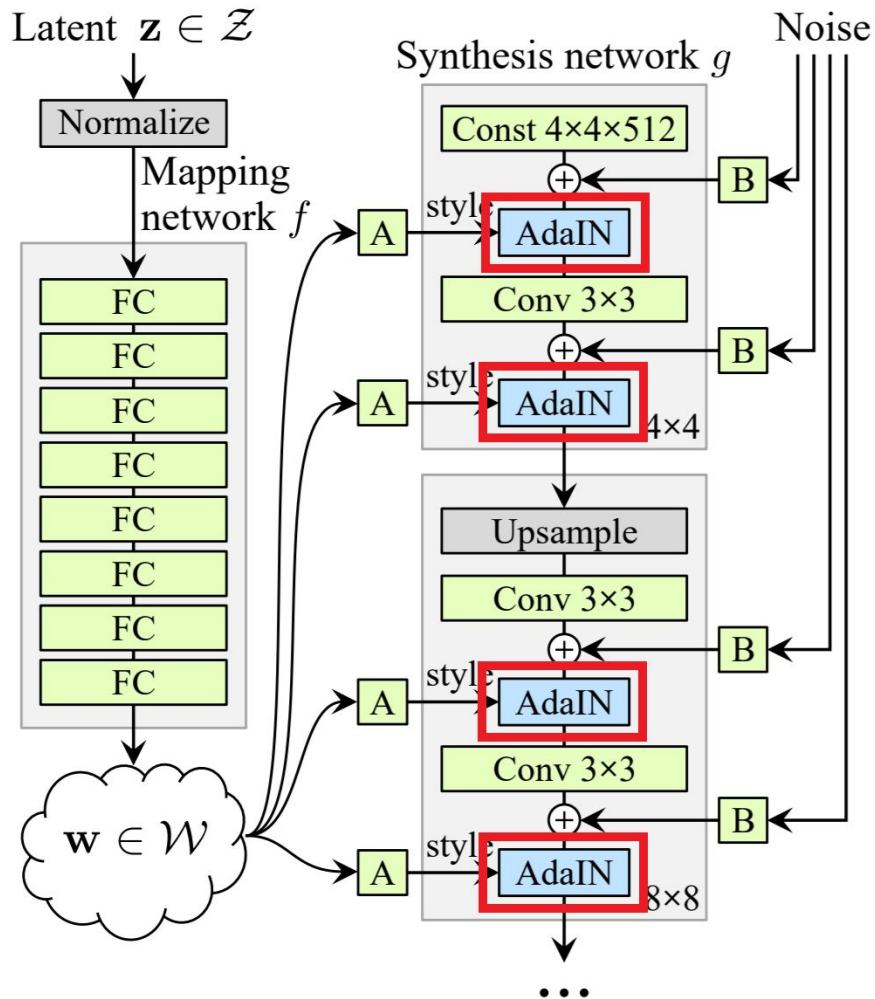


-Mapping network:  
наличие полносвязных слоёв  
помогает распутать пространство  $\mathcal{Z}$

-Аффинное преобразование с  
обучаемыми весами



- Mapping network:  
наличие полносвязных слоёв  
помогает распутать пространство  $\mathcal{Z}$
- Аффинное преобразование с  
обучаемыми весами
- Шум умножается на обучаемые  
веса и прибавляется к выходу  
свёрточных сетей.



-Mapping network:  
наличие полносвязных слоёв  
помогает распутать пространство  $\mathcal{Z}$

-Аффинное преобразование с  
обучаемыми весами

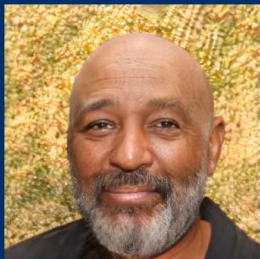
-Шум умножается на обучаемые  
веса и прибавляется к выходу  
свёрточных сетей.

-Adaptive Instance Normalization

$$\text{AdaIN}(\mathbf{x}_i, \mathbf{y}) = \mathbf{y}_{s,i} \frac{\mathbf{x}_i - \mu(\mathbf{x}_i)}{\sigma(\mathbf{x}_i)} + \mathbf{y}_{b,i}$$

# Смешение стилей

Coarse styles  
( $4^2 - 8^2$ )



Middle styles  
( $16^2 - 32^2$ )



Fine styles  
( $64^2 - 1024^2$ )





## Функция потерь

В качестве функции потерь для генератора используется non-saturating loss с  $R_1$ -регуляризацией:

$$L = -\mathbb{E}_z(\log(D(G(z)))) + \frac{\gamma}{2} \mathbb{E}_{p_D(x)} (\|\nabla D(x)\|)$$

# Метрики качества

Frechet Inception Distance:

$$\text{FID} = |\mu_D - \mu_G|^2 + \text{tr}(\Sigma_D + \Sigma_G - 2(\Sigma_D \Sigma_G)^{1/2}),$$

где  $\mu_D$  и  $\mu_G$  - выборочные средние, а  $\Sigma_D$  и  $\Sigma_G$  - матрицы выборочной ковариации векторных представлений реальных и сгенерированных изображений соответственно.

# Метрики качества

Perceptual path length:

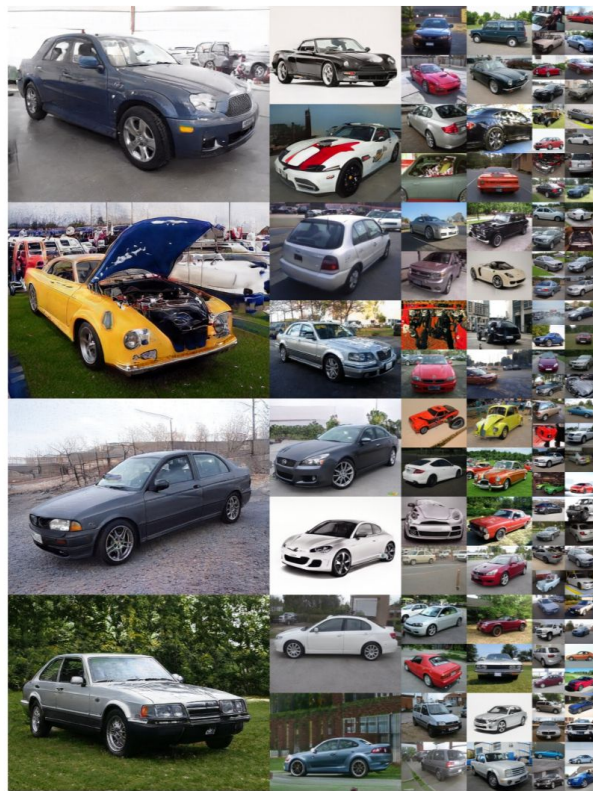
$$PPL = \mathbb{E} \left[ \frac{1}{\epsilon^2} d(G(slerp(z_1, z_2; t)), G(slerp(z_1, z_2; t + \epsilon))) \right],$$

где  $z_1, z_2 \sim P(z)$ ,  $t \sim U[0, 1]$ , *slerp* - сферическая интерполяция, *G* - генератор, *d* - Евклидово расстояние.

# Результаты

Method	CelebA-HQ	FFHQ
A Baseline Progressive GAN [30]	7.79	8.04
B + Tuning (incl. bilinear up/down)	6.11	5.25
C + Add mapping and styles	5.34	4.85
D + Remove traditional input	5.07	4.88
E + Add noise inputs	<b>5.06</b>	4.42
F + Mixing regularization	5.17	<b>4.40</b>

# Результаты

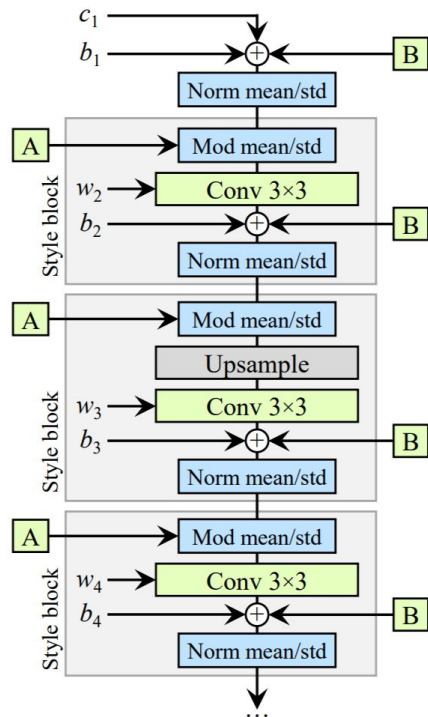


# Проблемы StyleGan

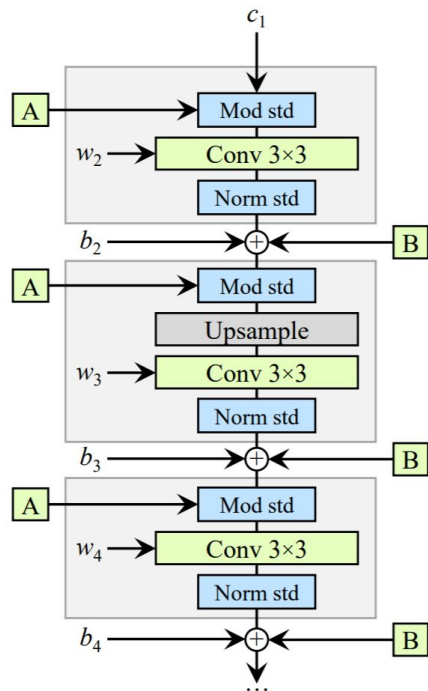


Характерной проблемой StyleGan является появление “пузырей” или “капель”, обычно находящихся на границе волос и заднего фона.

# StyleGan2



(b) StyleGAN (detailed)

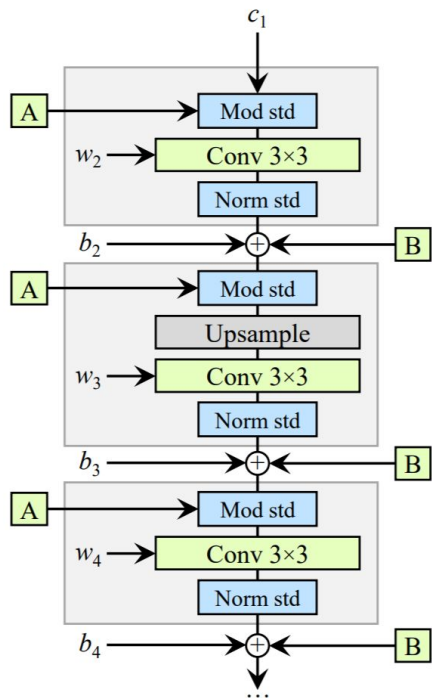


(c) Revised architecture

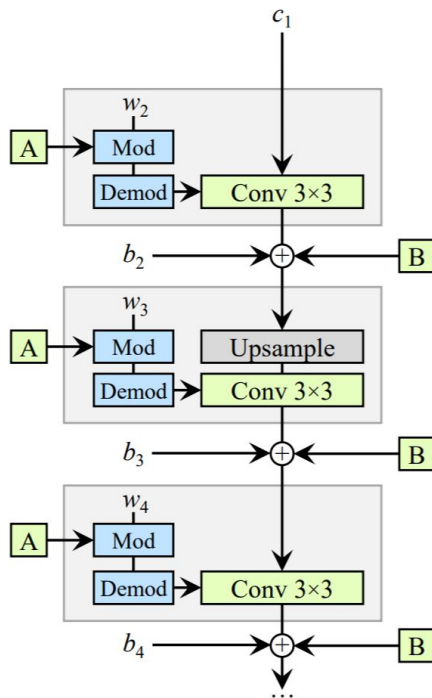
В StyleGan2 шум прибавляется к изображению после нормировки.

Также после этого оказывается достаточно приводить изображение к единичной дисперсии, оставляя ненулевое среднее.

# StyleGan2



(c) Revised architecture



(d) Weight demodulation

Вместо умножения картинки можно умножать веса самих свёрточных слоёв:

$$w'_{ijk} = s_i \cdot w_{ijk},$$

После этого можно также нормировать веса вместо нормировки данных:

$$w''_{ijk} = w'_{ijk} / \sqrt{\sum_{i,k} w'_{ijk}^2 + \epsilon},$$

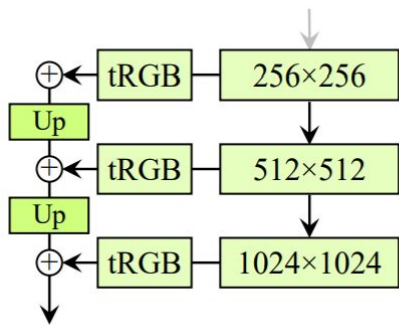


# Проблемы StyleGan

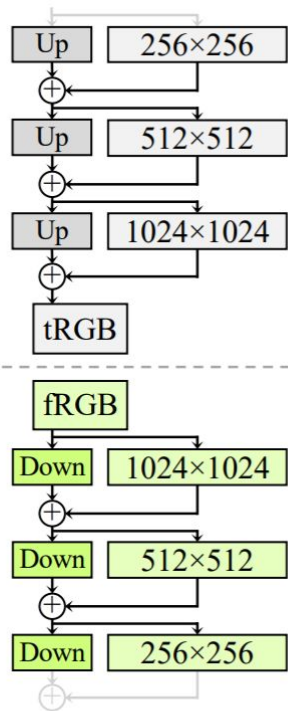


Ещё одной проблемой является фиксированное положение некоторых частей лица, таких как зубы или глаза.

# StyleGan2



(b) Input/output skips



(c) Residual nets

FFHQ	D original		D input skips		D residual	
	FID	PPL	FID	PPL	FID	PPL
G original	4.32	265	4.18	235	3.58	269
G output skips	4.33	169	3.77	127	<b>3.31</b>	<b>125</b>
G residual	4.35	203	3.96	229	3.79	243

LSUN Car	D original		D input skips		D residual	
	FID	PPL	FID	PPL	FID	PPL
G original	3.75	905	3.23	758	3.25	802
G output skips	3.77	544	3.86	<b>316</b>	3.19	471
G residual	3.93	981	3.40	667	<b>2.66</b>	645

# Регуляризация

Ленивая регуляризация:

Градиент регуляризационных слагаемых в функции потерь считается не каждый раз, тем самым уменьшая вычислительную сложность.

Path length регуляризация:

$$\mathbb{E}_{\mathbf{w}, \mathbf{y} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})} \left( \left\| \mathbf{J}_{\mathbf{w}}^T \mathbf{y} \right\|_2 - a \right)^2,$$

где  $\mathbf{w} \sim f(\mathbf{z})$ ,  $\mathbf{J}_{\mathbf{w}}$  - якобиан генератора,  $a$  - экспоненциальное скользящее среднее  $\|\mathbf{J}_{\mathbf{w}} \mathbf{y}\|$ .

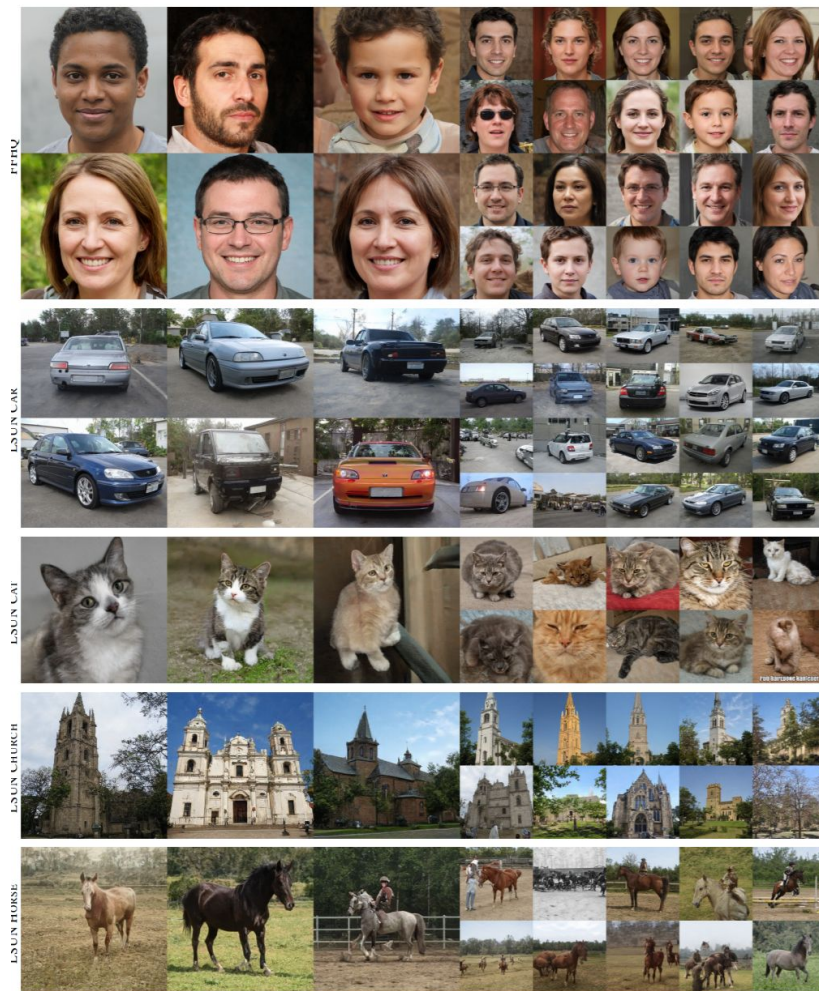
# Результаты

Configuration	FFHQ, 1024×1024				LSUN Car, 512×384			
	FID ↓	Path length ↓	Precision ↑	Recall ↑	FID ↓	Path length ↓	Precision ↑	Recall ↑
A Baseline StyleGAN [24]	4.40	212.1	<b>0.721</b>	0.399	3.27	1484.5	<b>0.701</b>	0.435
B + Weight demodulation	4.39	175.4	0.702	0.425	3.04	862.4	0.685	0.488
C + Lazy regularization	4.38	158.0	0.719	0.427	2.83	981.6	0.688	0.493
D + Path length regularization	4.34	<b>122.5</b>	0.715	0.418	3.43	651.2	0.697	0.452
E + No growing, new G & D arch.	3.31	124.5	0.705	0.449	3.19	471.2	0.690	0.454
F + Large networks (StyleGAN2)	<b>2.84</b>	145.0	0.689	<b>0.492</b>	<b>2.32</b>	<b>415.5</b>	0.678	<b>0.514</b>
Config A with large networks	3.98	199.2	0.716	0.422	–	–	–	–

Dataset	Resolution	StyleGAN (A)		StyleGAN2 (F)	
		FID	PPL	FID	PPL
LSUN CAR	512×384	3.27	1485	<b>2.32</b>	<b>416</b>
LSUN CAT	256×256	8.53	924	<b>6.93</b>	<b>439</b>
LSUN CHURCH	256×256	4.21	742	<b>3.86</b>	<b>342</b>
LSUN HORSE	256×256	3.83	1405	<b>3.43</b>	<b>338</b>



# Результаты



## Список источников

<https://arxiv.org/pdf/1812.04948.pdf>

<https://arxiv.org/pdf/1912.04958.pdf>

<https://towardsdatascience.com/explained-a-style-based-generator-architecture-for-gans-generating-and-tuning-realistic-6cb2be0f431>