

# StyleGAN

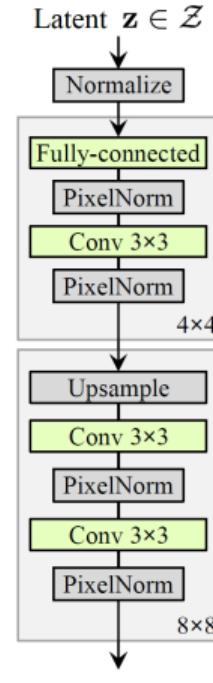
Шипилов Фома

ВШЭ

1 марта 2022 г.

# Мотивация StyleGAN

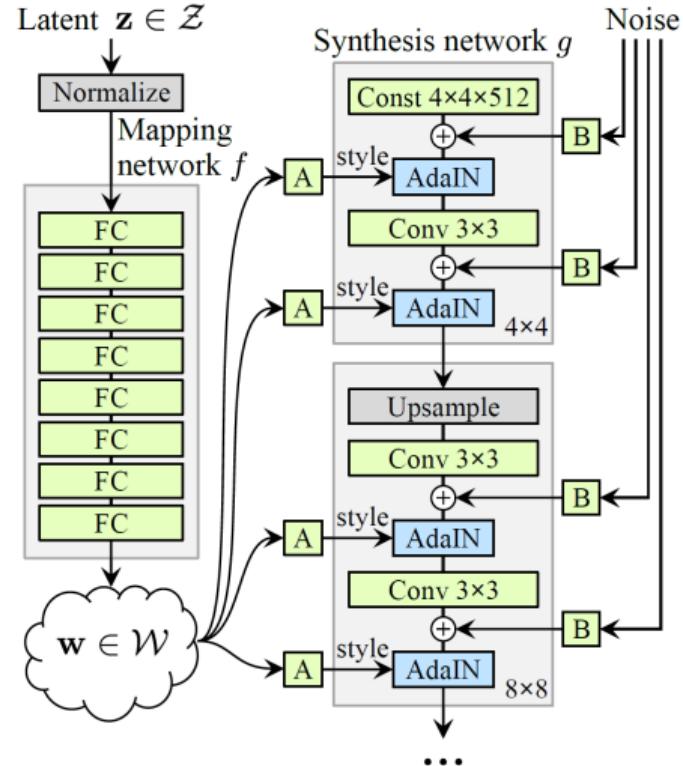
- Обычный GAN получает на вход случайный вектор из латентного пространства;
- Этот вектор одновременно вводит в модель элемент *случайности* и кодирует *признаковое описание* генерируемого сэмпла;
- Было бы логично разделить эти назначения.



(a) Traditional

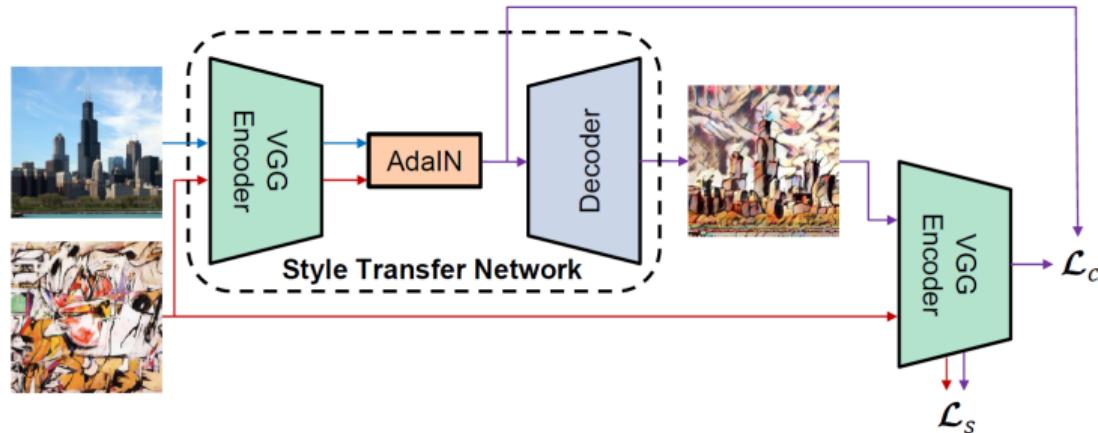
# Ключевые компоненты StyleGAN

- Латентное пространство;
- Многослойный перцептрон-кодировщик;
- Входы стиля (AdaIN);
- Входы шума;
- Дискриминатор.



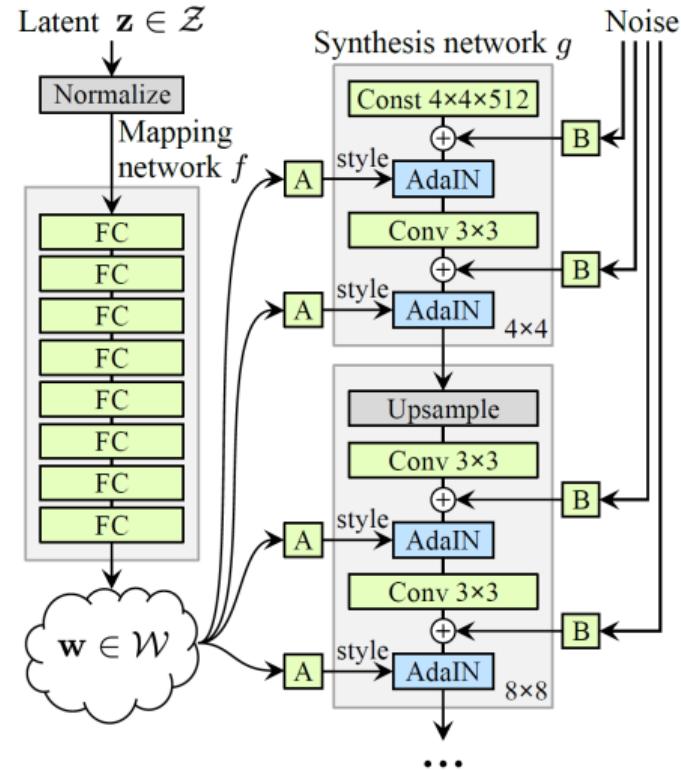
# Что такое стиль?

- Стиль — понятие из задачи *Style Transfer*;
- Численное представление высокоуровневых признаков стиля изображения;
- Один из способов представления стиля — относительные различия значений в картах признаков разных каналов.



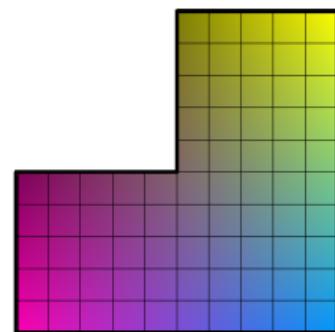
# Структура генератора

- Нет латентного вектора на входе, вместо него обучаемый константный тензор;
- Вместо этого стиль подается в модель в каждом слое через операцию **AdaIN** (*Adaptive Instance Normalization*);
- Перед каждым AdaIN ко входу добавляется шум;
- Билинейное масштабирование вместо транспонированной свертки между блоками.

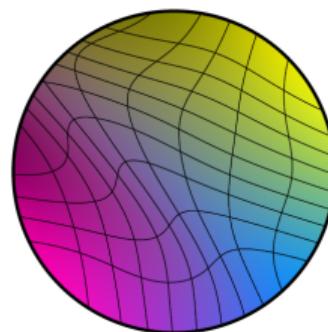


## Кодирование стиля

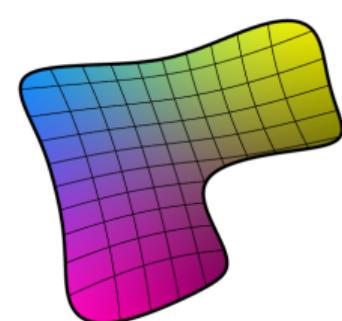
- В обычном GAN преобразование из латентного пространства в распределение картинок полностью выполняется в генераторе;
- Однако уже самые первые слои начинают конструировать изображение;
- В StyleGAN нелинейность выносится в перцептрон-кодировщик вне генератора, который преобразует вектор из латентного пространства в стиль.



(a) Distribution of  
features in training set



(b) Mapping from  
 $\mathcal{Z}$  to features



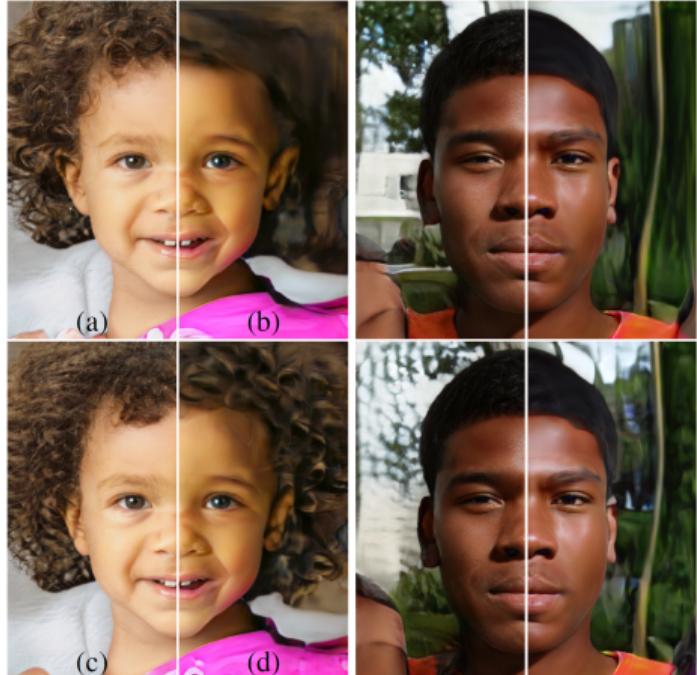
(c) Mapping from  
 $\mathcal{W}$  to features

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

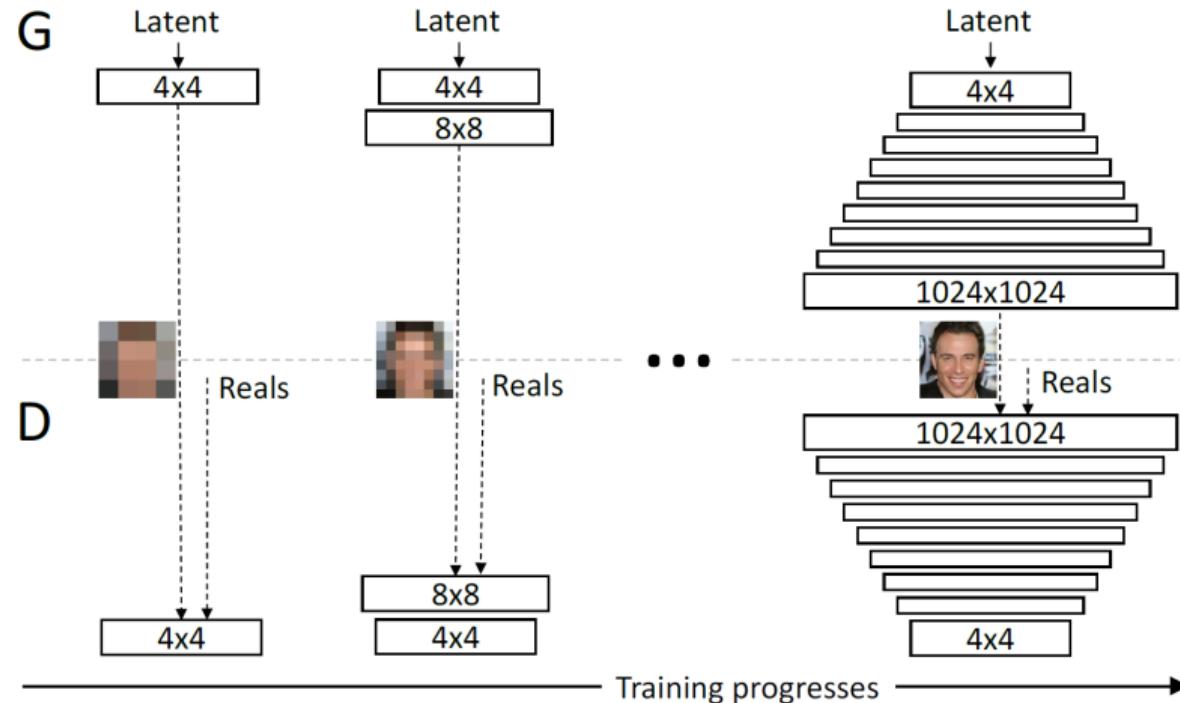
- $\mathbf{x}_i$  —  $i$ -й канал (двумерная карта признаков);
- Для каждого канала по два значения стиля  $\mathbf{y}_s, \mathbf{y}_b$  — масштабирование и сдвиг;
- $\mathbf{y} \in \mathbb{R}^{2 \times c}$  получается из стиля  $\mathbf{w} \in \mathbb{R}^{512}$  с помощью обучаемого аффинного преобразования (одно для всех слоев).

# Добавление шума

- Некоррелированный аддитивный гауссов шум;
- Для каждого канала обучаемый коэффициент масштабирования;
- Шум позволяет генерировать мелкие детали.



# Схема обучения

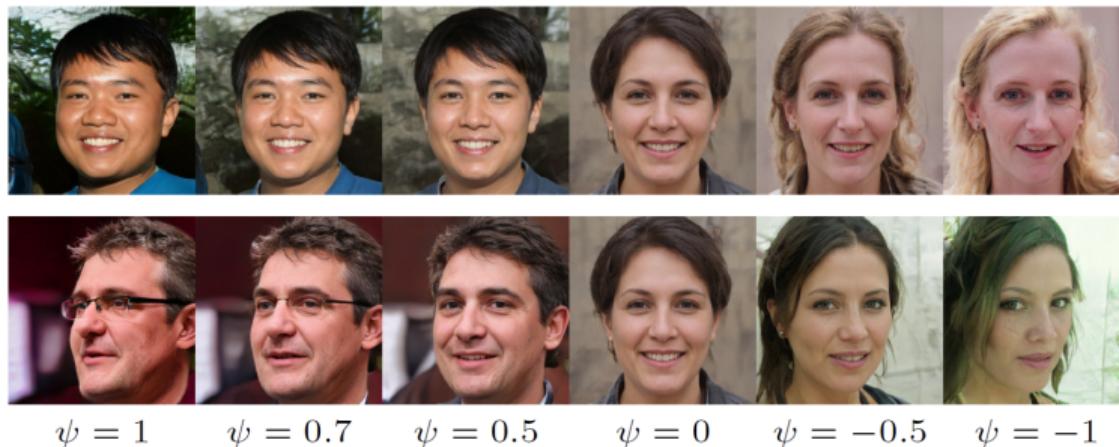


## Схема обучения

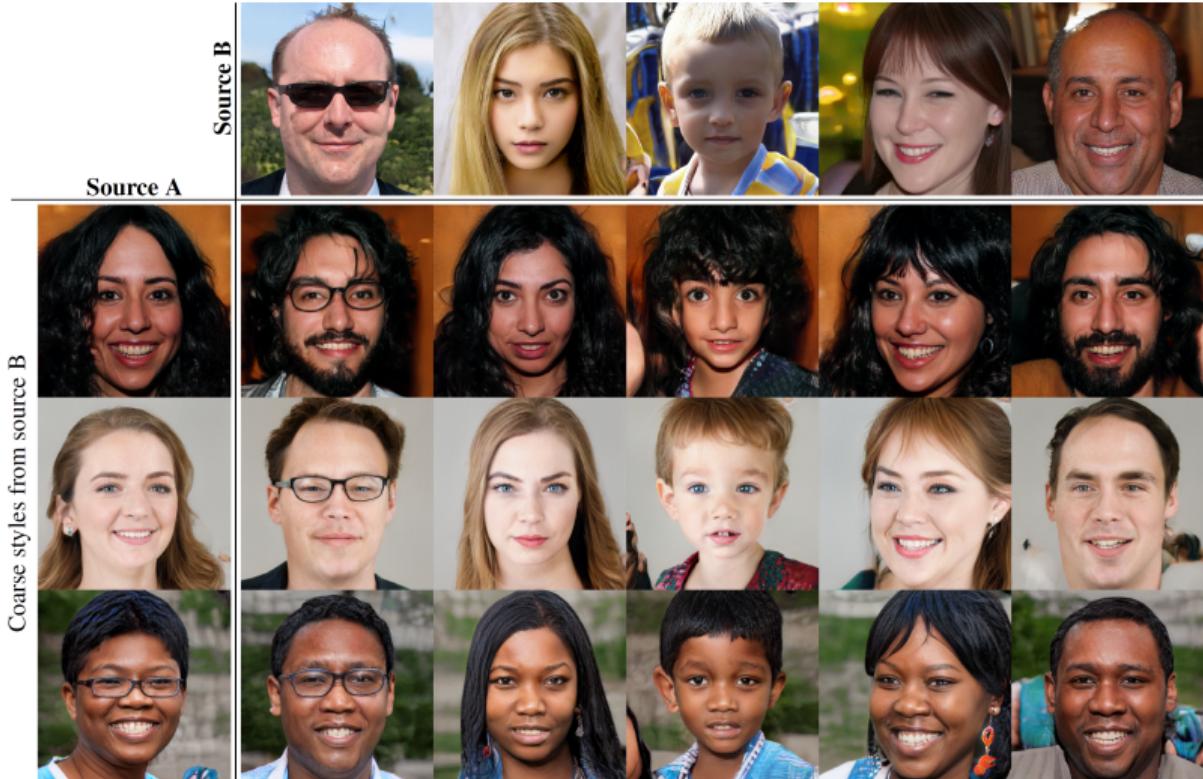
- Обучение идет не сразу всей модели end-to-end, а постепенно, то есть вначале обучается только первый слой генератора и последний дискриминатора, затем пристыковывается следующий и т.д. (Progressive GAN);
- 30% семплов обучаются на 2-х стилях вместо одного чтобы модель более полно их использовала (Mixing Regularization).

# Исследование латентного пространства

- После обучения в латентном пространстве нужно найти направления, отвечающие каким-то осмысленным признакам;
- Для этого можно отклассифицировать случайную выборку сгенерированных изображений и обучить на полученных классах SVM;
- Перпендикуляр к разделяющей гиперплоскости будет направлением.



# Результаты



# Результаты



## Список источников

- ① A Style-Based Generator Architecture for Generative Adversarial Networks
- ② Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization
- ③ Progressive Growing of GANs for Improved Quality, Stability, and Variation