

Alias-Free Generative Adversarial Networks (StyleGAN3)

Докладчик: Барановская Дарья

Рецензент: Латышев Александр

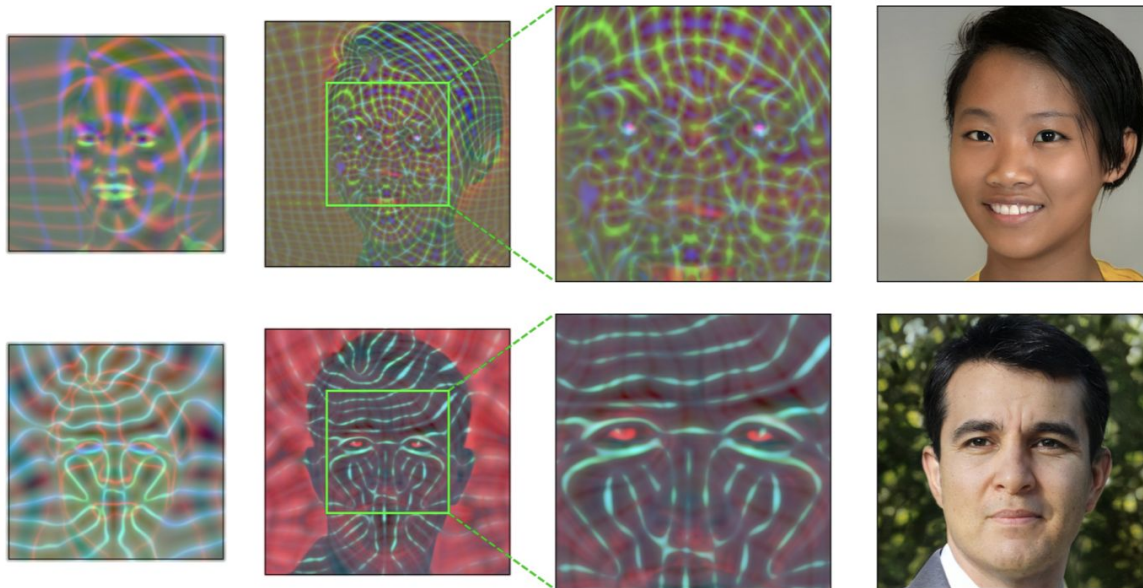
Практик-исследователь: Молодык Петр

Хакер: Щербинин Артём (умер от хакерства)

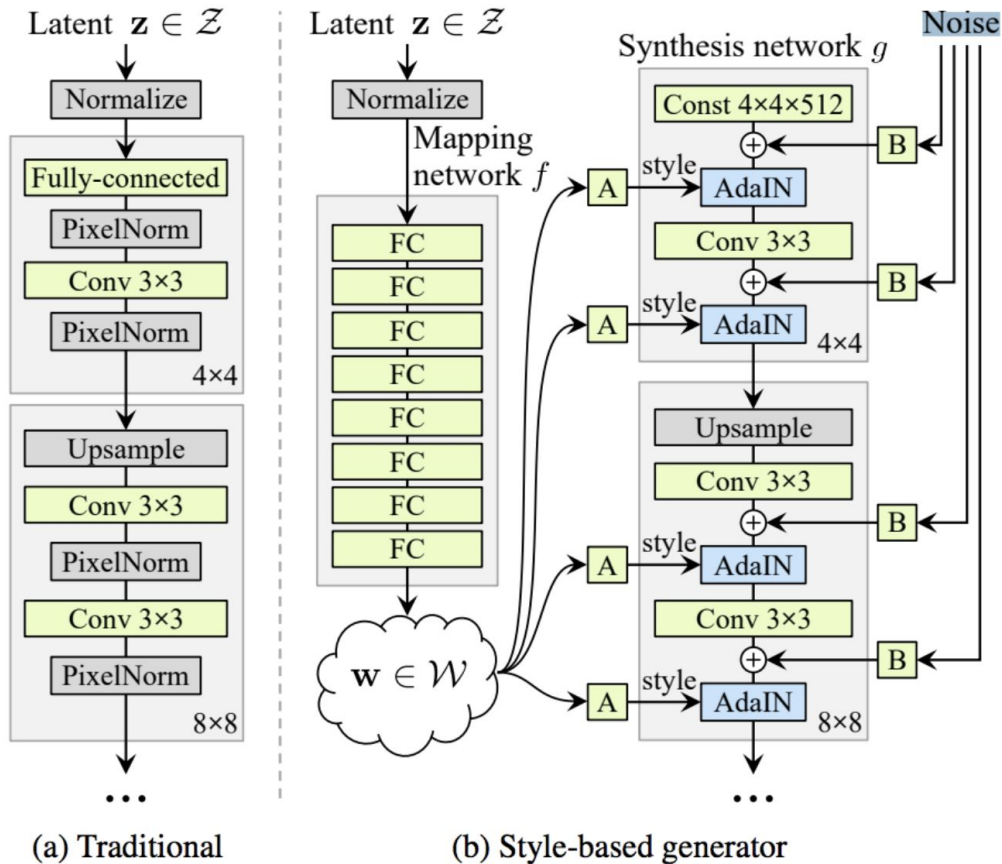
Alias-Free Generative Adversarial Networks

Tero Karras¹ Miika Aittala¹ Samuli Laine¹ Erik Härkönen^{2, 1} Janne Hellsten¹ Jaakko Lehtinen^{1, 2}
Timo Aila¹

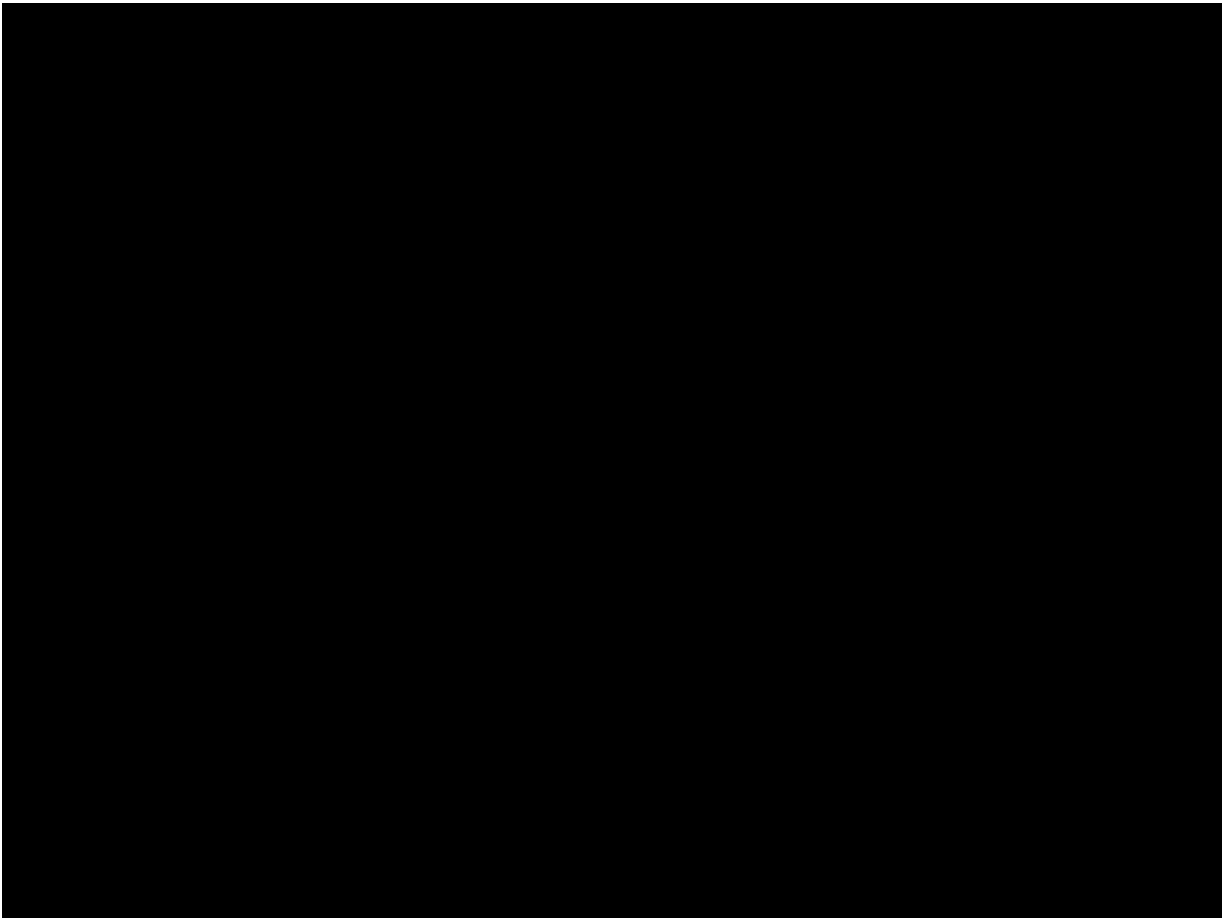
¹ NVIDIA ² Aalto University



StyleGAN, StyleGAN2



Проблемы StyleGAN2



Причины:

- 1) границы изображения
- 2) вводы шума для каждого пикселя
- 3) позиционное кодирование
- 4) алиасинг

Эквивариантность посредством интерпретации непрерывного сигнала

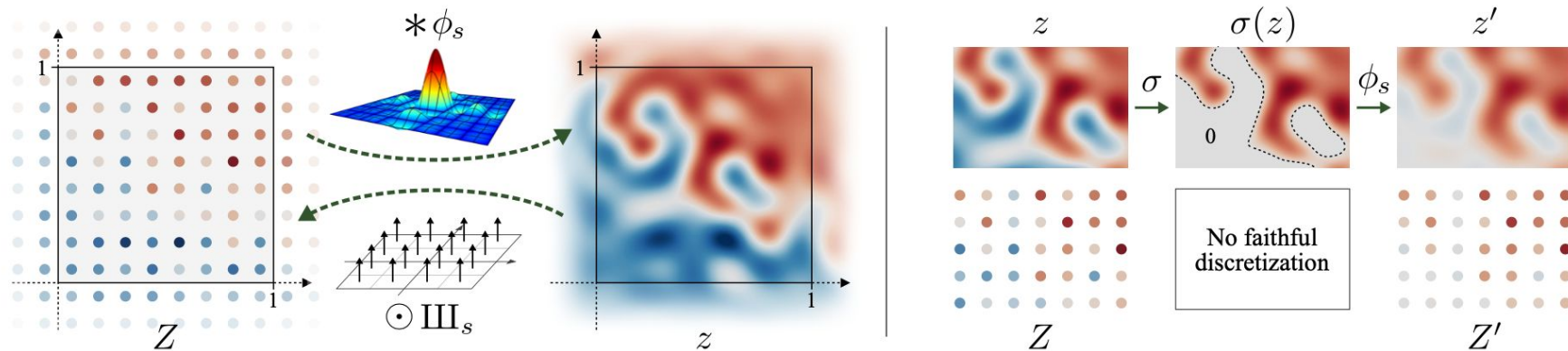
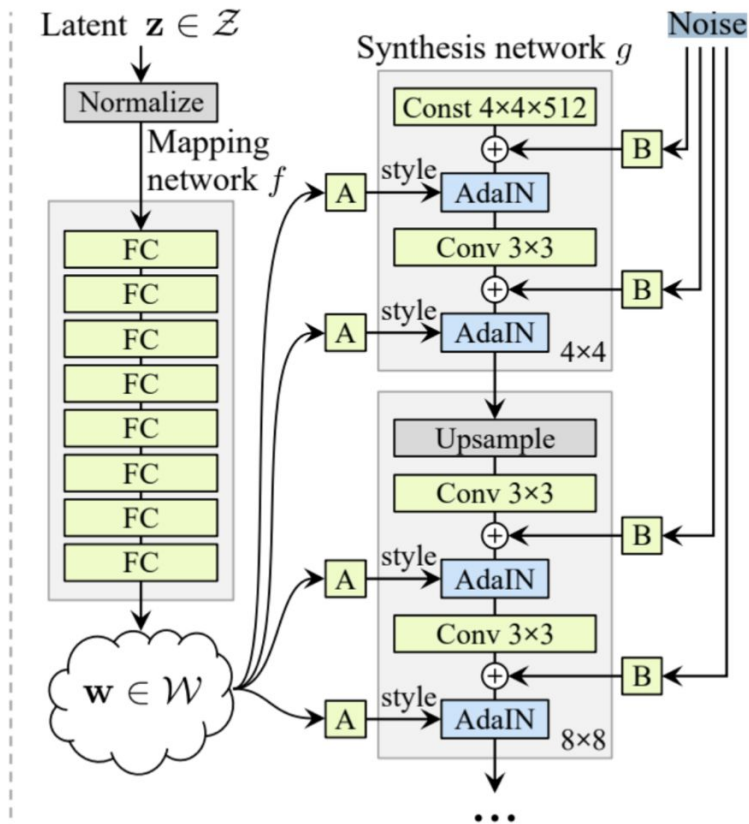


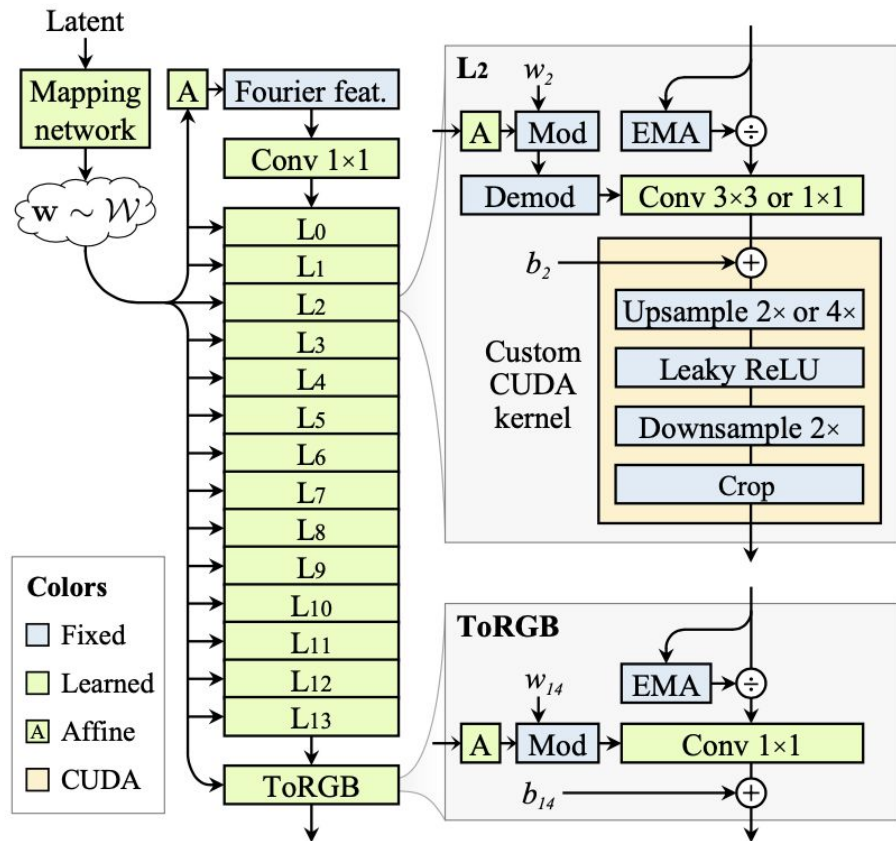
Figure 2: **Left:** Discrete representation Z and continuous representation z are related to each other via convolution with ideal interpolation filter ϕ_s and pointwise multiplication with Dirac comb $\textcircled{\cdot} \text{III}_s$. **Right:** Nonlinearity σ , ReLU in this example, may produce arbitrarily high frequencies in the continuous-domain $\sigma(z)$. Low-pass filtering via ϕ_s is necessary to ensure that Z' captures the result.

StyleGAN2



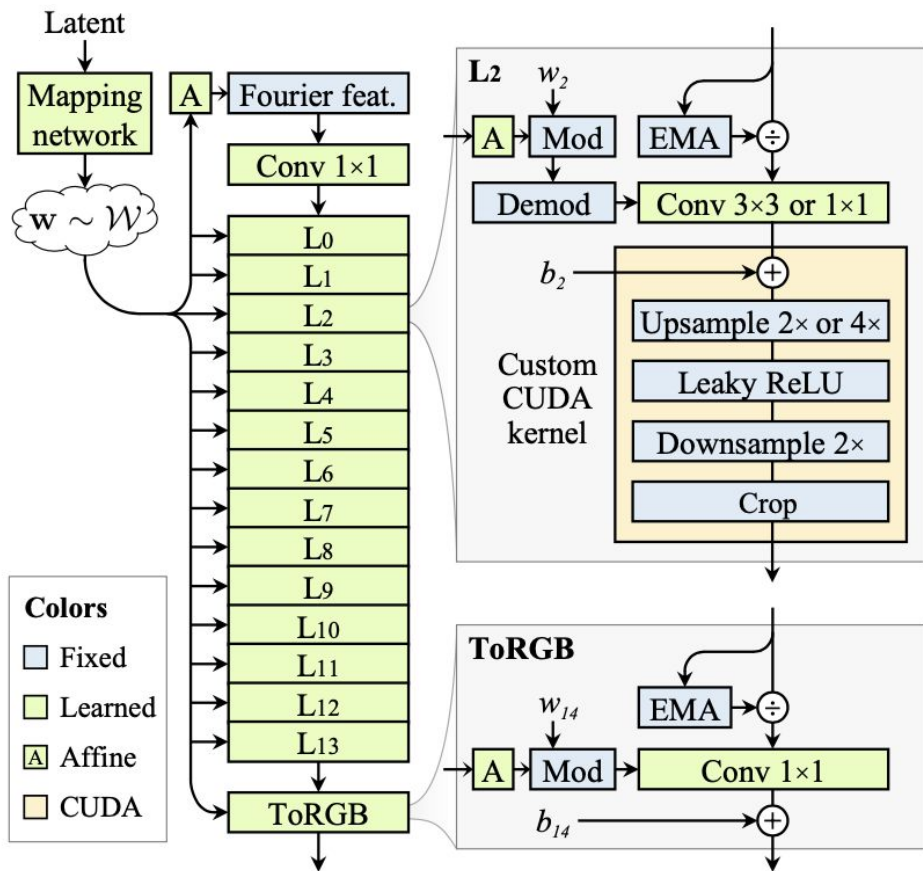
(b) Style-based generator

StyleGAN3



b) Our alias-free StyleGAN3 generator architecture

Архитектура StyleGAN3



- 1) Fourier feature на входе
- 2) Отказ от добавления шума на каждом слое
- 3) Упрощенный генератор
- 4) Границы и upsampling
- 5) Filtered nonlinearities
- 6) Non-critical sampling
- 7) Аффинные преобразования Fourier feature
- 8) Гибкие слои
- 9) Устойчивость к ротациям

b) Our alias-free StyleGAN3 generator architecture

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

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:**

$$\text{EQ-T} = 10 \cdot \log_{10} \left(I_{max}^2 / \mathbb{E}_{\mathbf{w} \sim \mathcal{W}, x \sim \mathcal{X}^2, p \sim \mathcal{V}, c \sim \mathcal{C}} \left[\left(\mathbf{g}(\mathbf{t}_x[z_0]; \mathbf{w})_c(p) - \mathbf{t}_x[\mathbf{g}(z_0; \mathbf{w})]_c(p) \right)^2 \right] \right)$$

PSNR метрика сравнения
инвариантности к трансформациям

Выводы

- 1) Такой же FID как у StyleGAN2
- 2) Инвариантен к ротации, хорош для анимации
- 3) Тяжелее в вычислительном отношении
- 4) Пока не очень пользуется популярностью

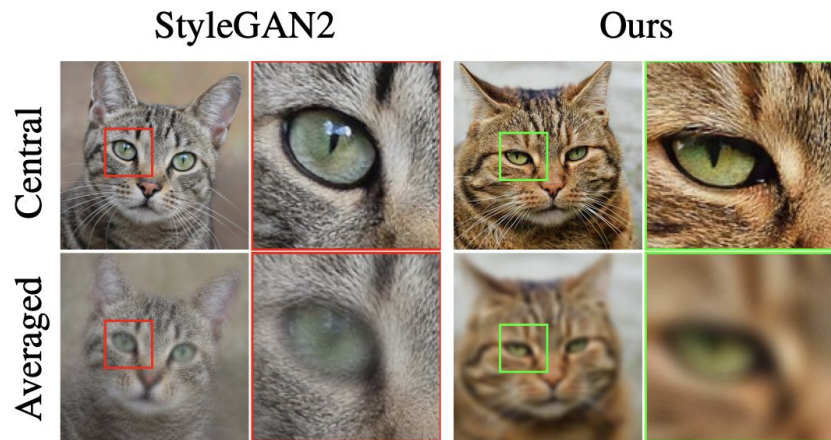


Figure 1: Examples of “texture sticking”. **Left**

Рецензент

Плюсы статьи

- Корректно написанный текст
- Полное теоретическое обоснование
- Большое число тестов
- Описаны сложности в обучении
- Репозиторий с кодом и готовой версией модели

Минусы статьи

- Не протестировано на архитектурах, кроме StyleGAN.
- Слишком сложное погружение в теоретическую часть.
- Легче подделать видео.

Другие рецензии

Есть небольшие замечания, но все согласны, что проделана большая работа в новом направлении.

Отсутствие тестов на других архитектурах.

Все рецензенты отмечают высокое качество изложения материала, некоторые приводят единичные замечания.

Все с положительной стороны смотрят на глубокое теоретическое погружение в проблему залипания структур.

Хорошее число тестов, подтверждающее работу алгоритма.

Практик-исследователь

История публикации

- Первая версия препринта опубликована 23 июня 2021
- Всего 4 версии препринта на архиве, последняя от 18 октября 2021, отличия между версиями по большей части косметические
- Принята на NeurIPS 2021 как oral
- Оценки 6, 10, 9, 8 на OpenReview

Авторы



Tero Karras



Timo Aila



Samuli Laine



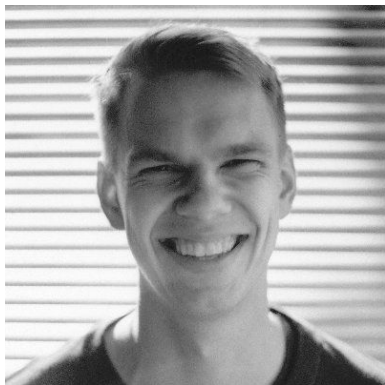
Miika Aittala



Janne Hellsten

Основная группа из NVIDIA, первые трое написали оригинальную статью про StyleGAN, другие двое присоединились к разработке StyleGAN2. Также вместе писали статью про рендеринг на конференцию по графике.

Авторы



Erik Härkönen



Jaakko Lehtinen

Jaakko - профессор в университете Аалто, был соавтором StyleGAN2, Erik - аспирант, в основном изучает контроль над работой ганов через латентные представления, самая цитируемая статья – про нахождение интерпретируемых латентных направлений в ганах.

Цитирования и связанные статьи

- В препринте *Toward Spatially Unbiased Generative Models* от авторов из университета Сеула и Samsung исследуется альтернативный вариант удаления неявной позиционной информации из латентных пространств ганов с помощью введения явного кодирования.
- В работе TransGAN статья упоминается вскользь в контексте выбора правильных методов апсемплинга.
- Упоминается в нескольких более мелких работах о инвариантности сверточных сетей к мелким сдвигам.

Применение и будущее

- Статья помогает бороться с “залипанием текстур”, поэтому построенные на стайлгане анимации и 3д-изображения выглядят более плавно и реалистично, поэтому можно попробовать применять новую версию стайлгана в этих областях
- Одна из важных задач на будущее – понять, насколько методы, протестированные на StyleGAN2, окажутся применимыми для других архитектур ганов.

Источники:

[1] <https://nvlabs.github.io/stylegan3/>