

GANs & Style Transfer



Алексей Сильвестров

R&D инженер в Samsung AI Center Moscow. Выпускник ВМиК МГУ.

LaMa: Large Mask Inpainting with Fourier Convolutions

Roman Suvorov
Arsenii Ashukha

Elizaveta Logacheva
Aleksii Silvestrov
Kiwoong Park

Anton Mashikhin
Naejin Kong
Victor Lempitsky

Anastasia Remizova
Harshith Goka

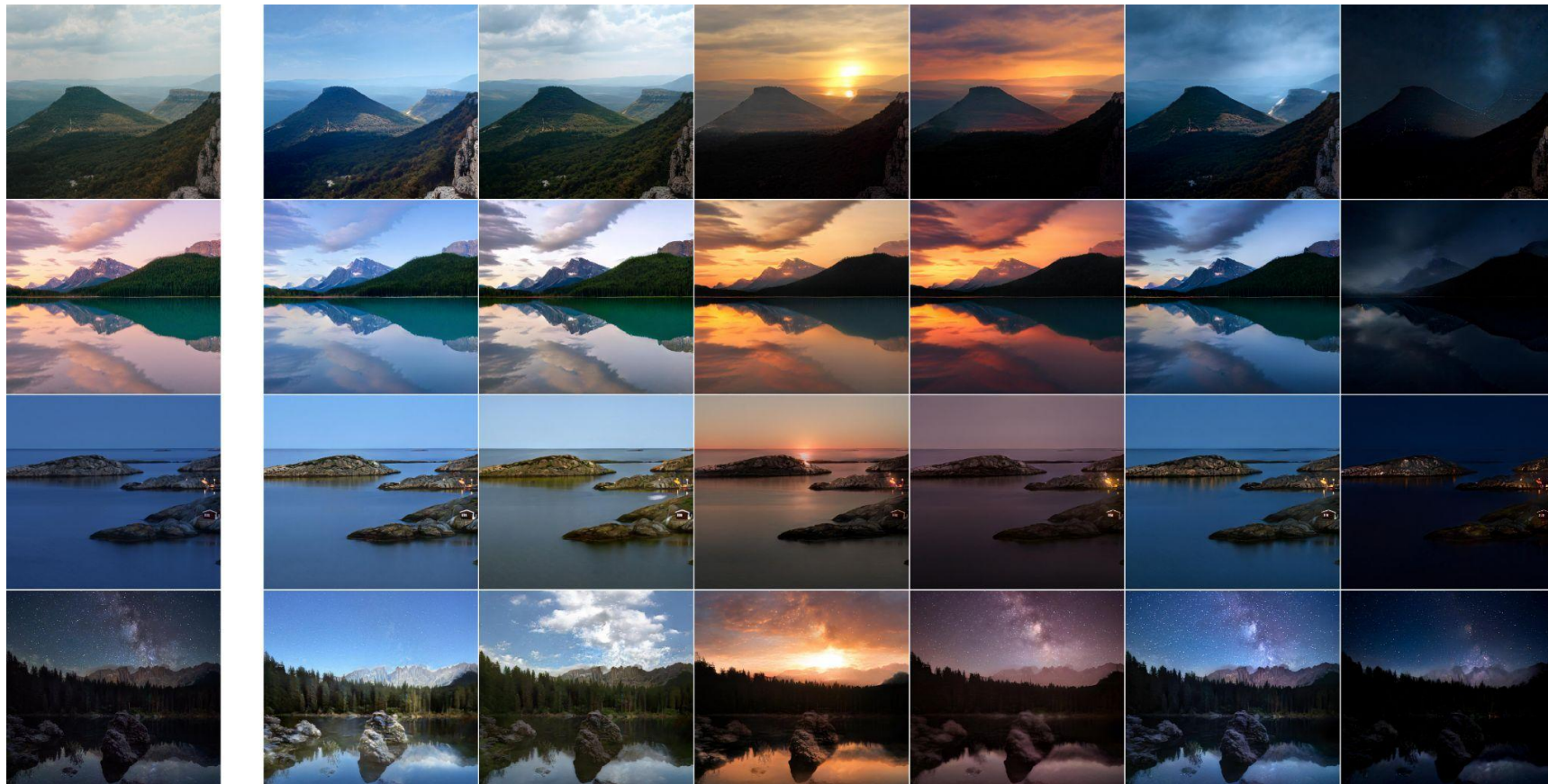


☆ Star 4.9k

advimman.github.io/lama-project

High Resolution Daytime Translation without Domain Labels

I. Anokhin^{1*}, P. Solovlev^{1*}, D. Korzhenkov^{1*}, A. Kharlamov^{1*},
T. Khakhulin^{1,3}, A. Silvestrov¹, S. Nikolenko^{2,1}, V. Lempitsky^{1,3}, G. Sterkin¹



Нейронные сети и обработка текста

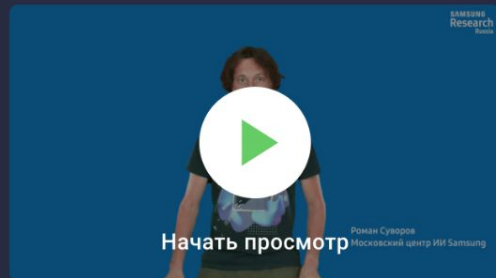
Современные методы автоматической обработки текста – это поиск по смыслу, машинный перевод, чат-боты, построение баз знаний... Как к этому подступиться? Больше практики! Авторы курса, эксперты Центра ИИ Samsung, доступным языком рассказывают, как начать работать с текстами при помощи нейросетей.



3-5 часов в
неделю



Сертификат
Stepik



★★★★★ 4.8

27 668 учащихся

[72 отзыва](#)



Алексей Сильвестров

R&D инженер в Samsung AI Center Moscow. Выпускник ВМиК МГУ.

<https://stepik.org/course/54098>

GANs & Aiva.ai



Aiva.ai

- Объявление в VK: видели 100 человек, написали 5, на интервью пришли 2
- Фаундер выступил на TED
- Новый домен: нет сеток ни для чего
- Гудфеллоу и Хуангу не показалось бредом генерить картинку с нотами



GANs

2017-2018

Wasserstein GAN

Martin Arjovsky¹, Soumith Chintala², and Léon Bottou^{1,2}

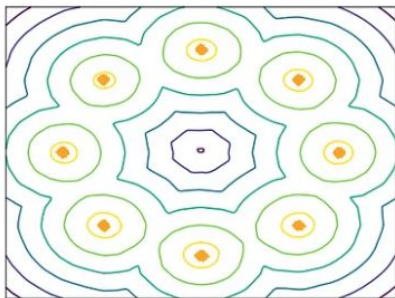
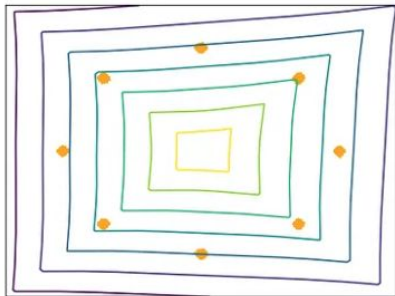
**The Cramer Distance as a Solution to Biased
Wasserstein Gradients**

Fisher GAN

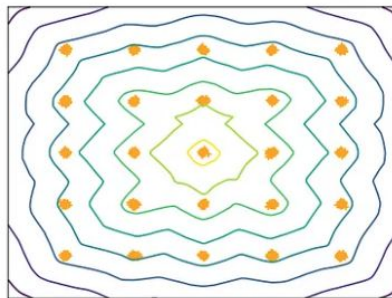
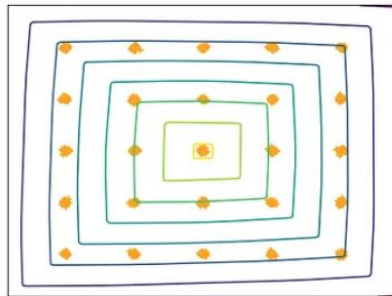
**PROGRESSIVE GROWING OF GANs FOR IMPROVED
QUALITY, STABILITY, AND VARIATION**

Wasserstein GAN

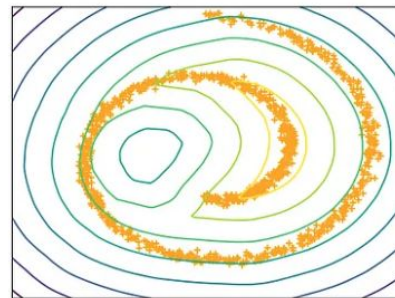
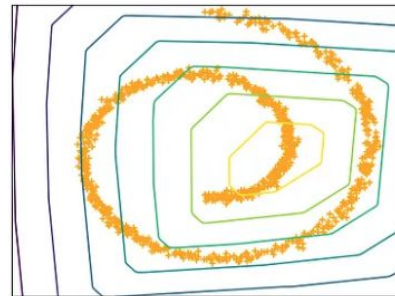
8 Gaussians



25 Gaussians



Swiss Roll



Wasserstein GAN

что-то про функан и оптимизацию

Optimal transport dual

- **Primal:**

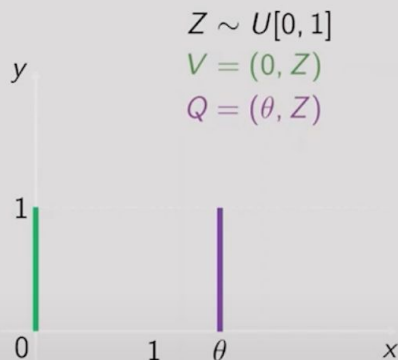
$$T(P, Q) = \inf_{\Gamma \in \mathcal{P}(x \sim P, y \sim Q)} \mathbb{E}_{(x, y) \sim \Gamma} [c(x, y)]$$

- **Dual (Wasserstein-1 metric):**

$$T(P, Q) = W_1(P, Q) = \sup_{\|f\|_L \leq 1} \mathbb{E}_{x \sim P} f(x) - \mathbb{E}_{x \sim Q} f(x)$$

Wasserstein GAN

Optimal transport vs f-Divergence

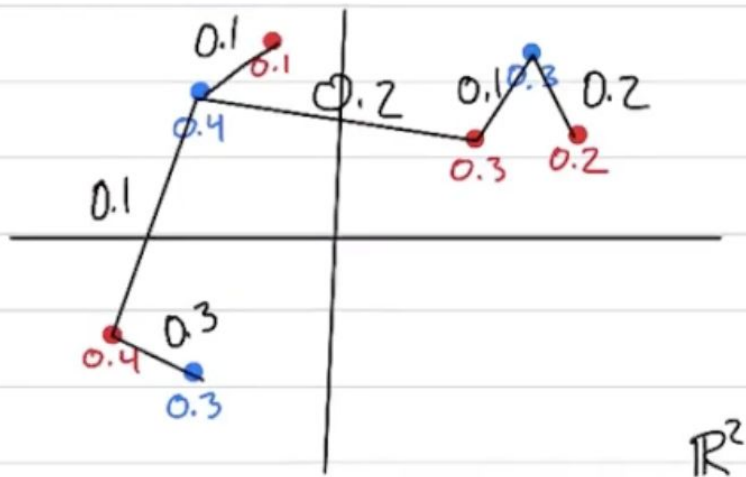


- $W_1(P, Q) = \theta$

- $JS(P\|Q) = \begin{cases} \log(2), & \theta \neq 0 \\ 0, & \theta = 0 \end{cases}$

- $KL(P\|Q) = \begin{cases} \infty, & \theta \neq 0 \\ 0, & \theta = 0 \end{cases}$

Earth Mover Distance



$$\mathbb{E}_{(x,y) \sim \gamma} [\|x - y\|]$$

		x_1	x_2	x_3	x_4	
		0.4	0.1	0.3	0.2	
y_1	0.3	0.3	0	0	0	Transport plan
y_2	0.4	0.1	0.1	0.2	0	
y_3	0.3	0	0	0.1	0.2	

Cramer GAN

критика праотца

- 1. The authors claim the energy distance kernel might be a better choice for GAN training. I am not convinced of the argument: theory-wise....
- 2. The Critic (discriminator) is not correct: Unfortunately, the paper makes a problematic approximation, which causes the critic not to correctly compare and match the generator and reference distributions.

Cramer GAN

критика праотца

- 1. The authors claim the energy distance kernel might be a better choice for GAN training. I am not convinced of the argument: theory-wise....
- 2. The Critic (discriminator) is not correct: Unfortunately, the paper makes a problematic approximation, which causes the critic not to correctly compare and match the generator and reference distributions.
- В итоге, если учесть комментарии Греттона, мы получим MMD-GAN
- Статью в итоге не приняли

Which Training Methods for GANs do actually Converge?

Method	Local convergence (a.c. case)	Local convergence (general case)
unregularized (Goodfellow et al., 2014)	✓	✗
WGAN (Arjovsky et al., 2017)	✗	✗
WGAN-GP (Gulrajani et al., 2017)	✗	✗
DRAGAN (Kodali et al., 2017)	✓	✗
Instance noise (Sønderby et al., 2016)	✓	✓
ConOpt (Mescheder et al., 2017)	✓	✓
Gradient penalties (Roth et al., 2017)	✓	✓
Gradient penalty on real data only	✓	✓
Gradient penalty on fake data only	✓	✓

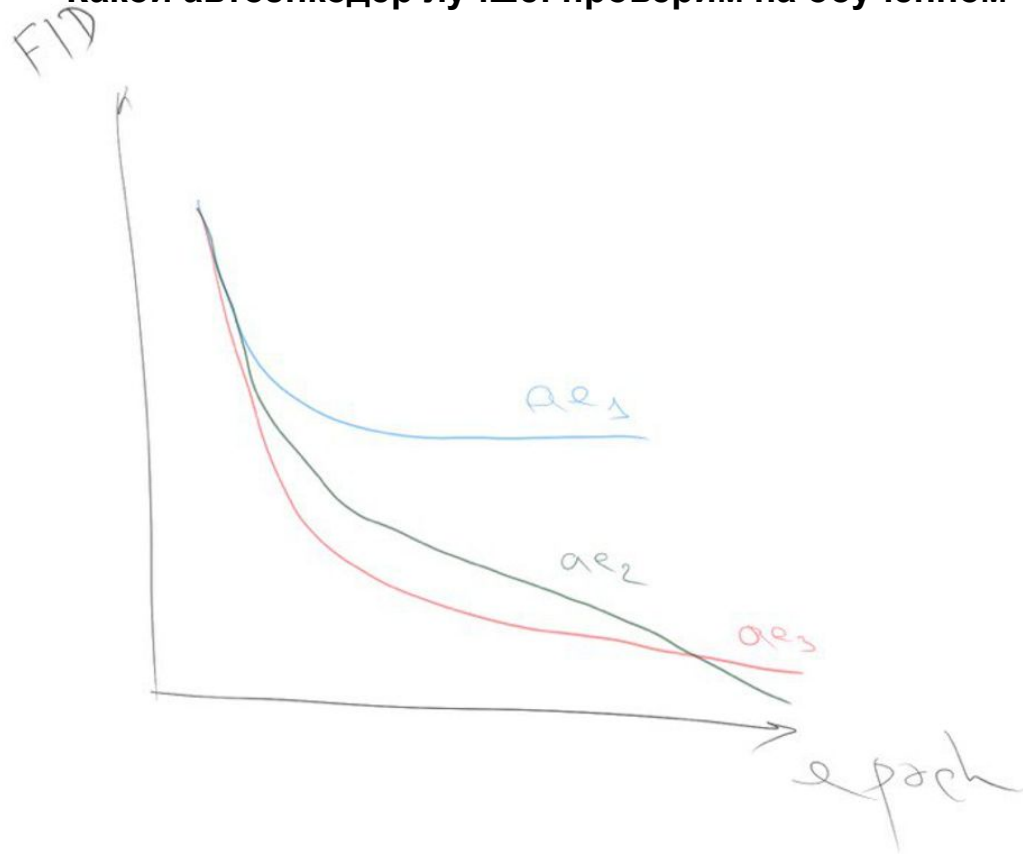
Frechet Inception Distance

$$FID = \|\mu_r - \mu_g\|^2 + \text{Tr}(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2})$$

- Сегодня стала одной из дефолтных метрик:
- ◆ Inception V3 в основе
 - ◆ Лица (?)
 - ◆ Любые картинки природы, города и интерьера

Frechet Domain Distance

Какой автоэнкодер лучше: проверим на обученном GAN



Frechet <...> Distance

- Посмотрите на Kernel Inception Distance: в диффузных статьях он снова мелькает
- Из любопытства посмотрите на Frechet Audio Distance:
<https://arxiv.org/pdf/1812.08466.pdf>

Style Transfer

2015-2020



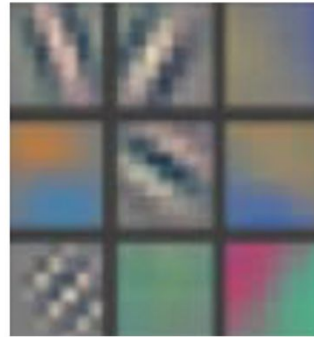
Фильтры

обученные

Patches from Input Image



Layer 1 Structure



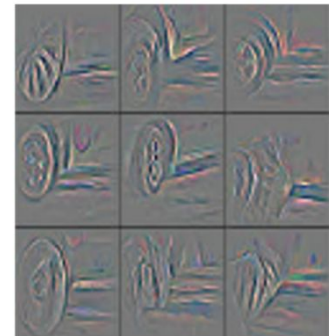
The first hidden layer learns to identify basic structural elements such as edges and color blobs

Patches from Input Image

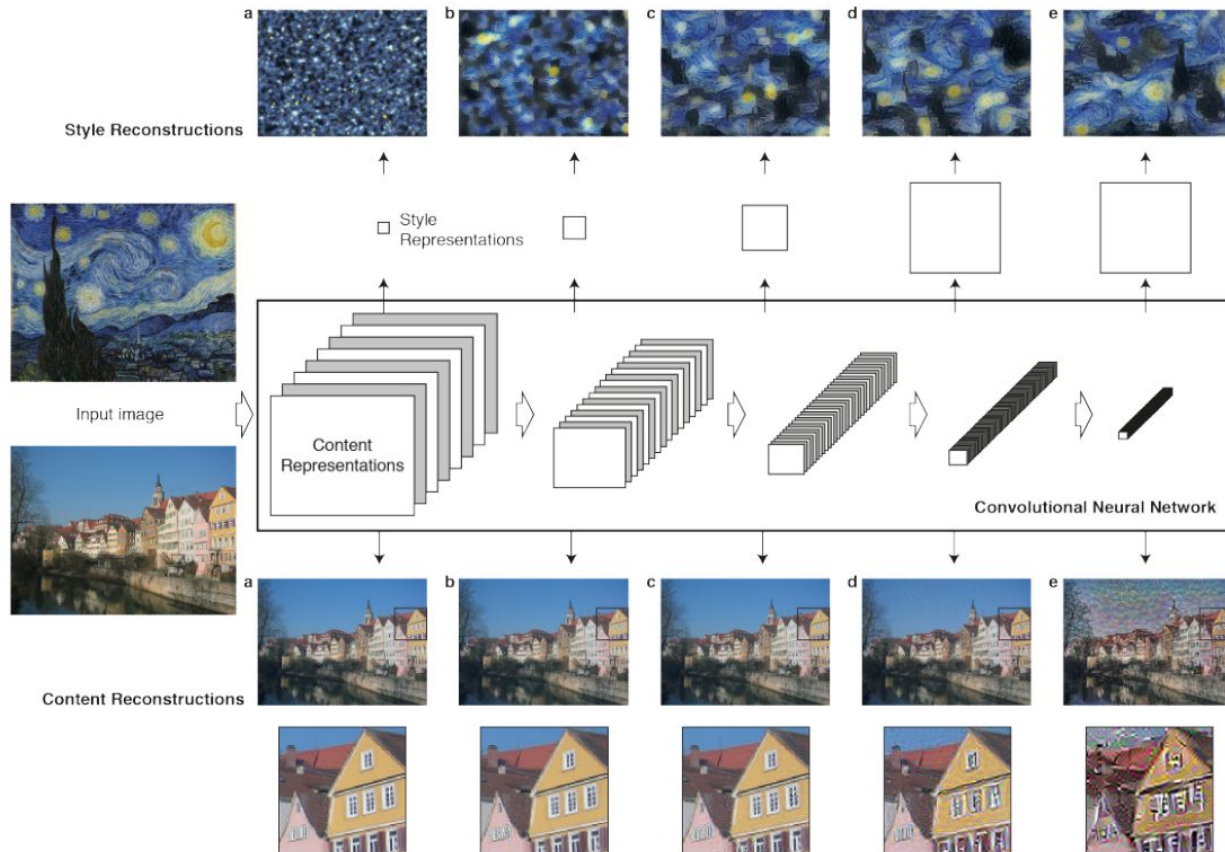


CNNs learn hierarchical structure after several layers

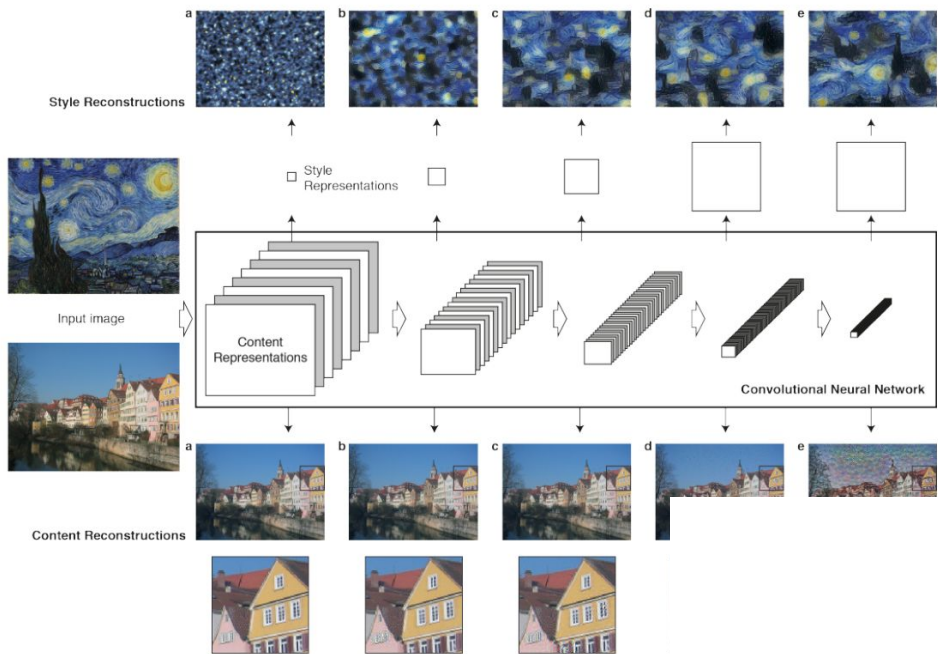
Layer 5 Structure



A Neural Algorithm of Artistic Style



A Neural Algorithm of Artistic Style

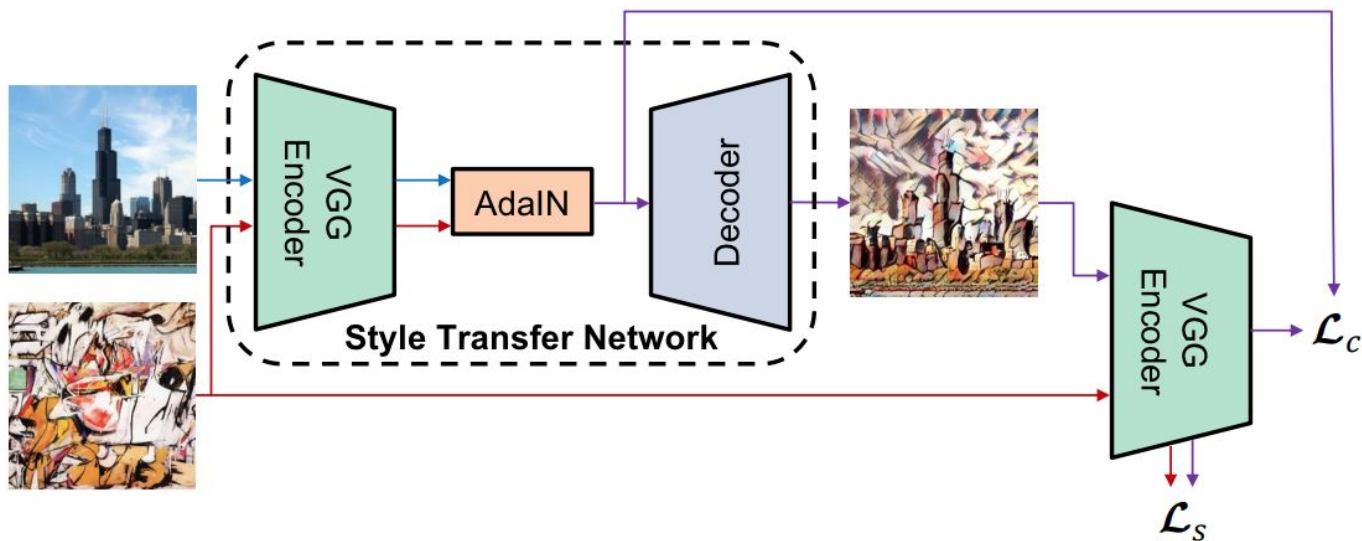


$$\mathcal{L}_{total}(\vec{p}, \vec{a}, \vec{x}) = \alpha \mathcal{L}_{content}(\vec{p}, \vec{x}) + \beta \mathcal{L}_{style}(\vec{a}, \vec{x})$$

$$G_{ij}^l(\mathbf{x}) = \langle F_i^l(\mathbf{x}), F_j^l(\mathbf{x}) \rangle$$

$$|G(\{v_1, \dots, v_n\})| = \begin{vmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle & \dots & \langle v_1, v_n \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \dots & \langle v_2, v_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_n, v_1 \rangle & \langle v_n, v_2 \rangle & \dots & \langle v_n, v_n \rangle \end{vmatrix}.$$

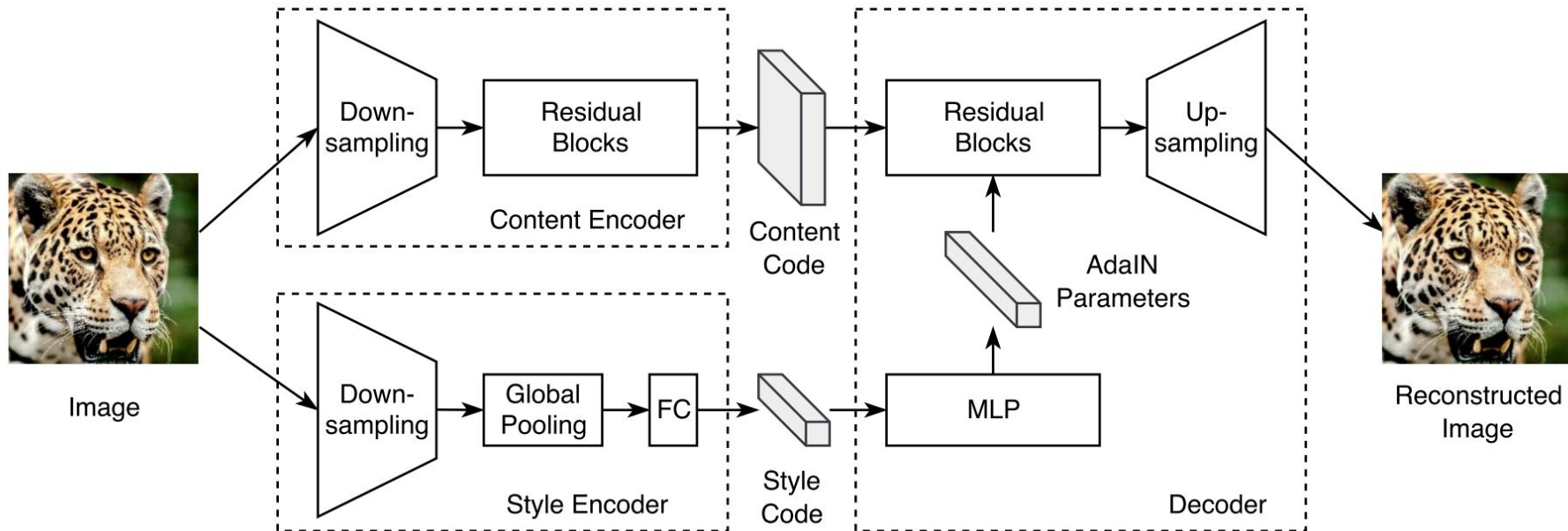
InstanceNorm -> CondIN -> AdaIN



Instead, it adaptively computes the affine parameters from the style input:

$$\text{AdaIN}(x, y) = \sigma(y) \left(\frac{x - \mu(x)}{\sigma(x)} \right) + \mu(y) \quad (8)$$

M-UNIT



$$\mathcal{L}_{\text{GAN}}^{x_2} = \mathbb{E}_{c_1 \sim p(c_1), s_2 \sim q(s_2)} [\log(1 - D_2(G_2(c_1, s_2)))] + \mathbb{E}_{x_2 \sim p(x_2)} [\log D_2(x_2)]$$

HiDT

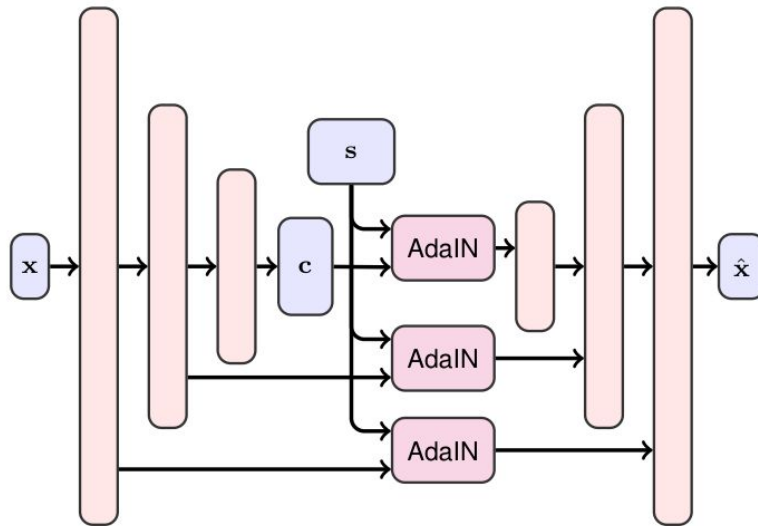


Figure 2: Diagram of the Adaptive U-Net architecture: an encoder-decoder network with dense skip-connections and content-style decomposition (c, s).

