

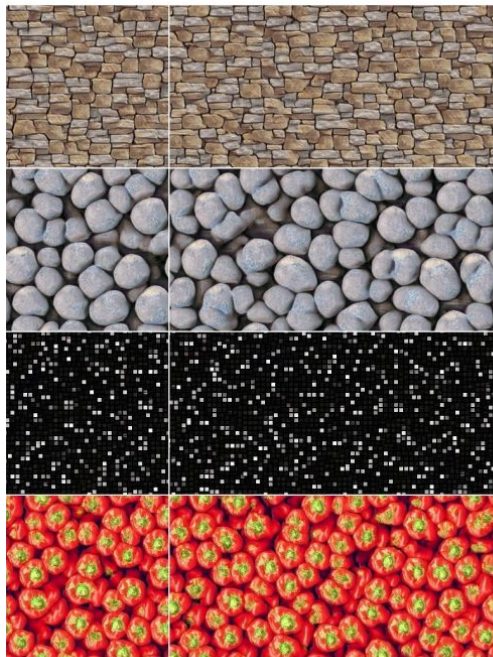
Neural Style Transfer

Гайдамашко Даниил, БПМИ-161

Постановка задачи

Texture generation

$$\mathbf{x} \sim p(\mathbf{x}|\mathbf{x}_0)$$



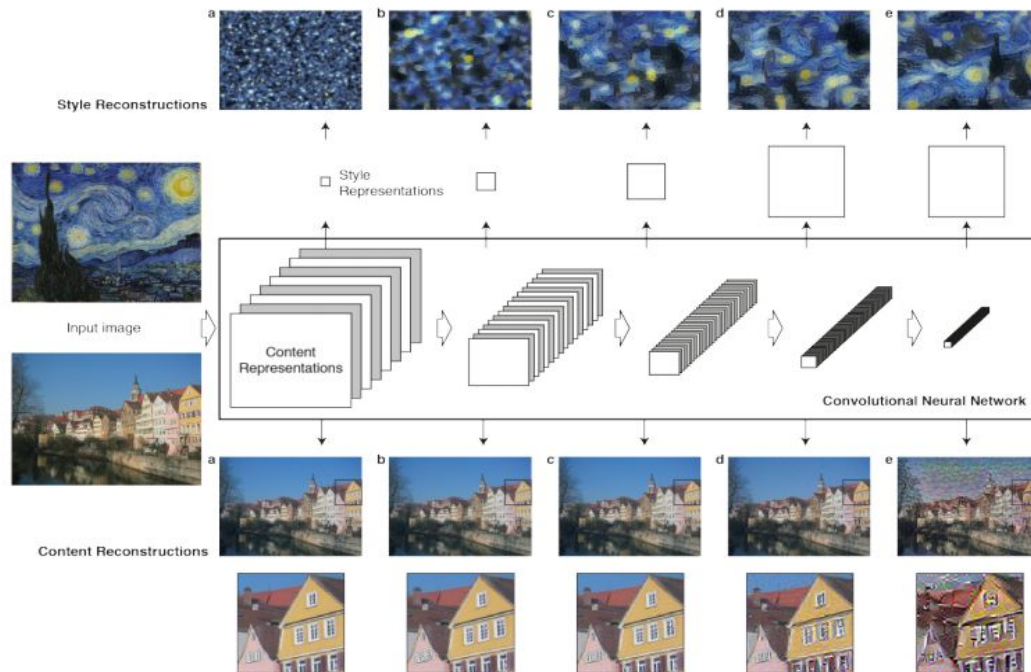
Style transferring

$$\mathbf{x} \sim p(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1)$$



Gatys et al.

- Дескриптивный подход;
- Показано, что сверточная сеть способна разделить стиль и содержание;
- Алгоритм затратный по времени и памяти;



Gatys et al.

- CNN на основе VGG, 16 Conv и 5 AvgPooling
- Восстанавливаем картинку на основе текущего слоя
- Строим новое признаковое пространство

Восстановление содержания

$$\mathcal{L}_{content}(\vec{p}, \vec{x}, l) = \frac{1}{2} \sum_{i,j} (F_{ij}^l - P_{ij}^l)^2$$

$$\frac{\partial \mathcal{L}_{content}}{\partial F_{ij}^l} = \begin{cases} (F^l - P^l)_{ij} & \text{if } F_{ij}^l > 0 \\ 0 & \text{if } F_{ij}^l < 0 . \end{cases}$$

Восстановление стиля

$$G_{ij}^l = \sum_k F_{ik}^l F_{jk}^l.$$

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{ij}^l - A_{ij}^l)^2$$

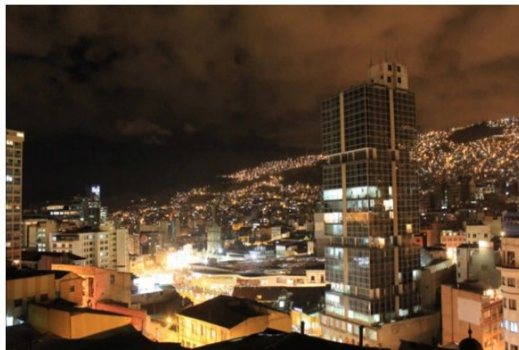
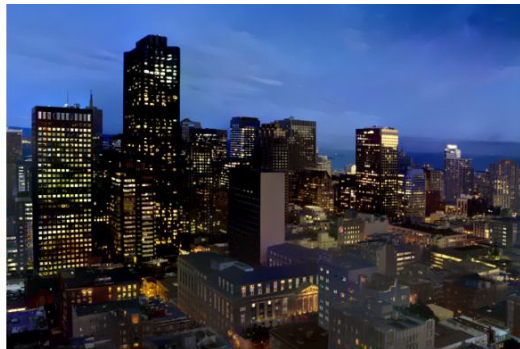
$$\mathcal{L}_{style}(\vec{a}, \vec{x}) = \sum_{l=0}^L w_l E_l$$

Trade-off

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



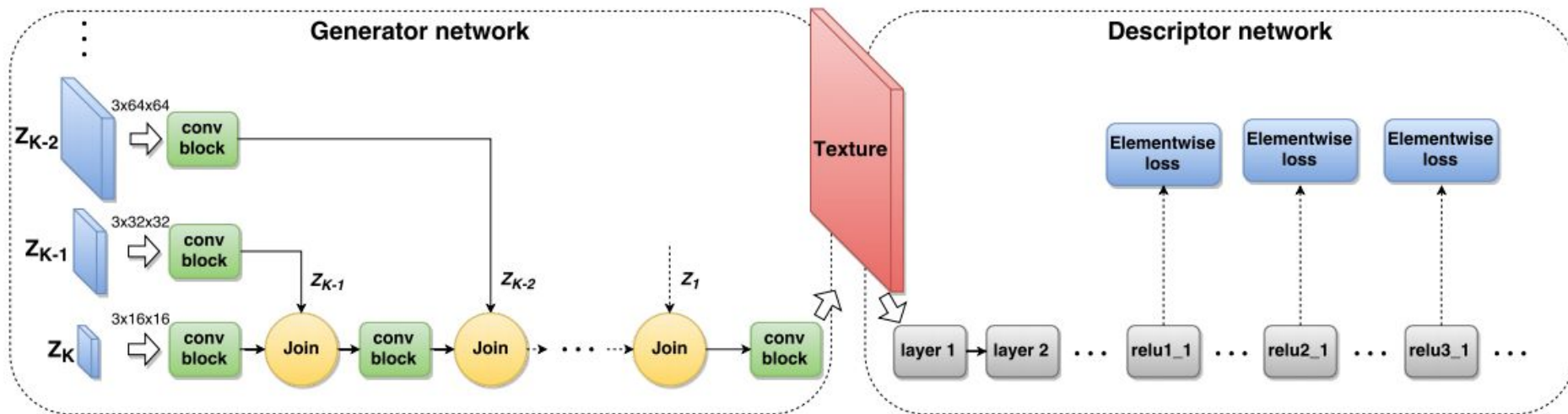
Deep Photo Style Transfer



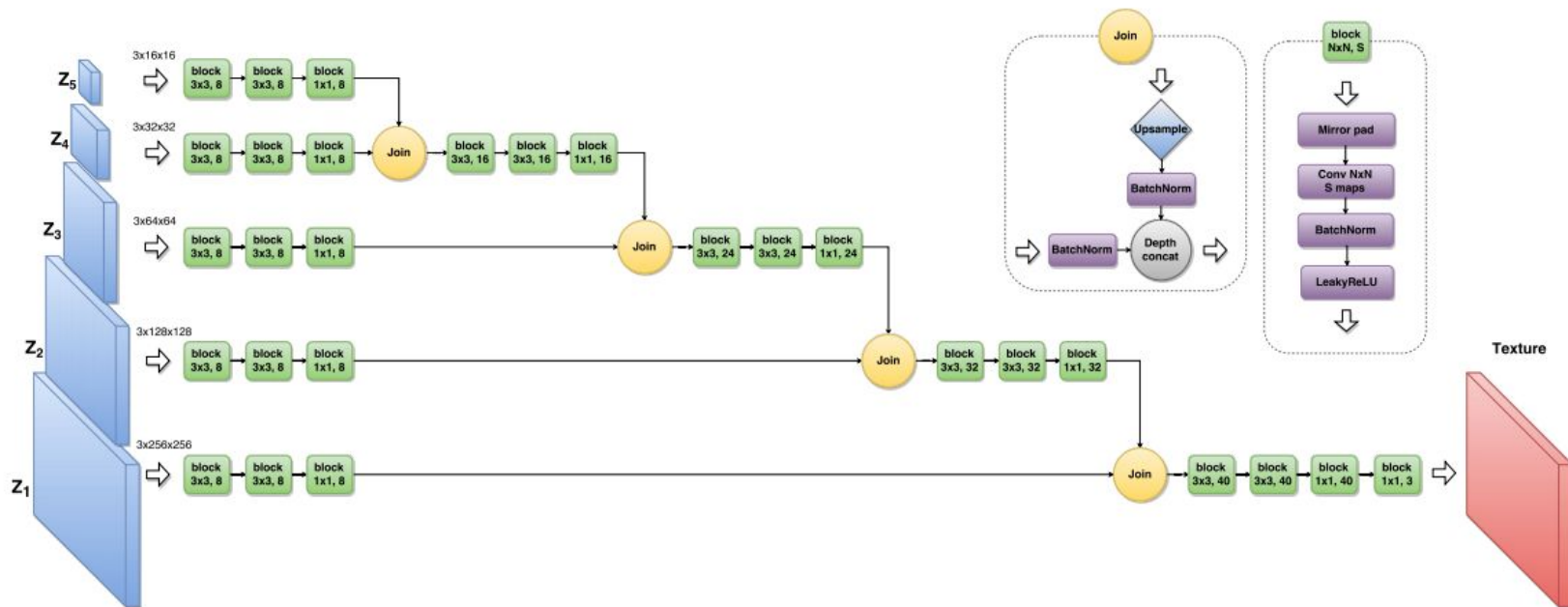
Ulyanov et al.

- Генеративный подход;
- Архитектура Texture Network;
- Существенное улучшение потребления времени и памяти при том же качестве;

Архитектура



Generator



Обучение

- Вычисление генератора
- Вычисление дескриптора
- Вычисление loss
- Backprop
- Обновление параметров

Texture generation

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

$$\mathcal{L}_T(\mathbf{x}; \mathbf{x}_0) = \sum_{l \in L_T} \|G^l(\mathbf{x}) - G^l(\mathbf{x}_0)\|_2^2.$$

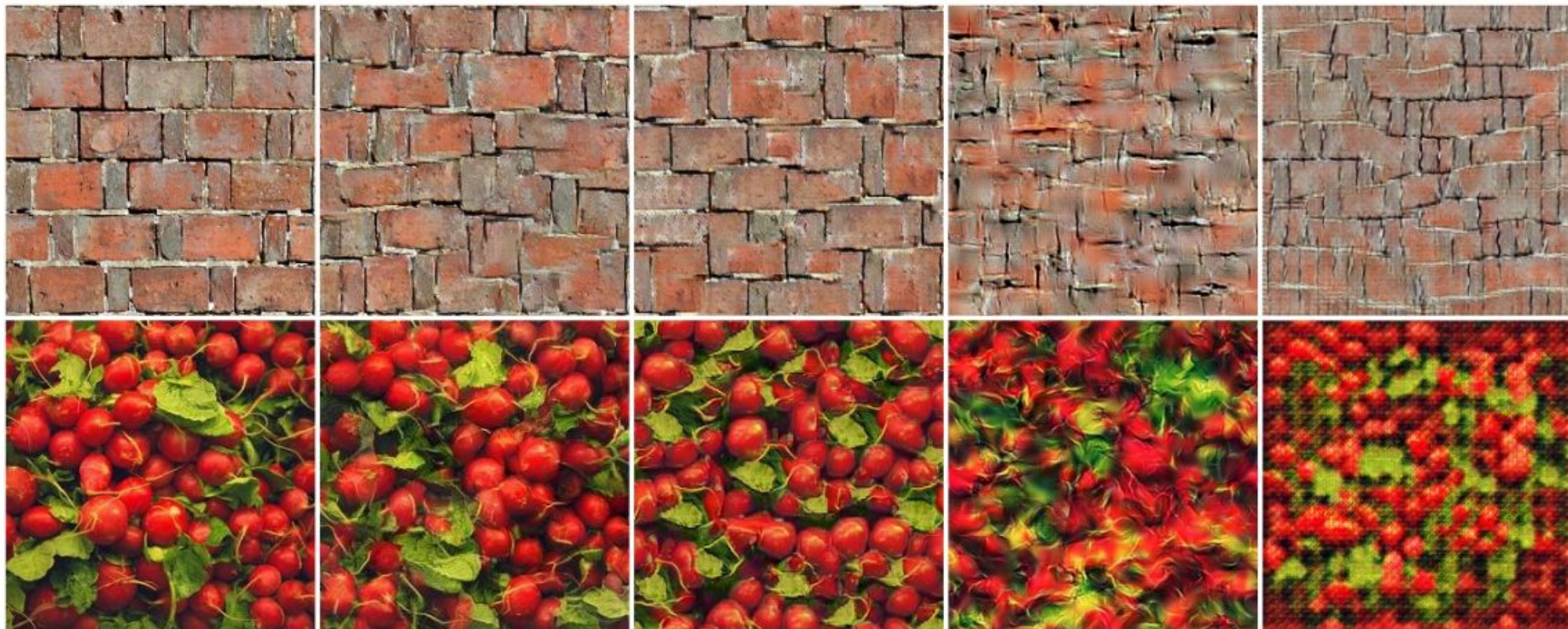
$$\mathcal{L}_C(\mathbf{x}; \mathbf{y}) = \sum_{l \in L_C} \sum_{i=1}^{N_l} \|F_i^l(\mathbf{x}) - F_i^l(\mathbf{y})\|_2^2$$

Style transfer

$$\theta_{\mathbf{x}_0} = \operatorname{argmin}_{\theta} E_{\mathbf{z} \sim \mathcal{Z}} [\mathcal{L}_T(\mathbf{g}(\mathbf{z}; \theta), \mathbf{x}_0)]$$

$$\theta_{\mathbf{x}_0} = \operatorname{argmin}_{\theta} E_{\mathbf{z} \sim \mathcal{Z}; \mathbf{y} \sim \mathcal{Y}} [\mathcal{L}_T(\mathbf{g}(\mathbf{y}, \mathbf{z}; \theta), \mathbf{x}_0) + \alpha \mathcal{L}_C(\mathbf{g}(\mathbf{y}, \mathbf{z}; \theta), \mathbf{y})].$$

Сравнение методов



Input

Gatys et al.

Texture nets (ours)

Portilla, Simoncelli

DCGAN

Сравнение методов



Content

Texture nets (ours)

Gatys et al.

Style

Выводы

Gatys et al.:

- Пионеры в нейросетевом подходе к решению задачи;
- Использование выходов CNN как описательных статистик;

Ulyanov et al.:

- Усовершенствование нейросетевого подхода;
- Генеративная архитектура с генератором и дескриптором;

ИСТОЧНИКИ

- <https://arxiv.org/abs/1508.06576> - Gatys and al.
- <https://arxiv.org/abs/1603.03417> - Ulyanov and al.
- <https://habr.com/ru/post/402665/>