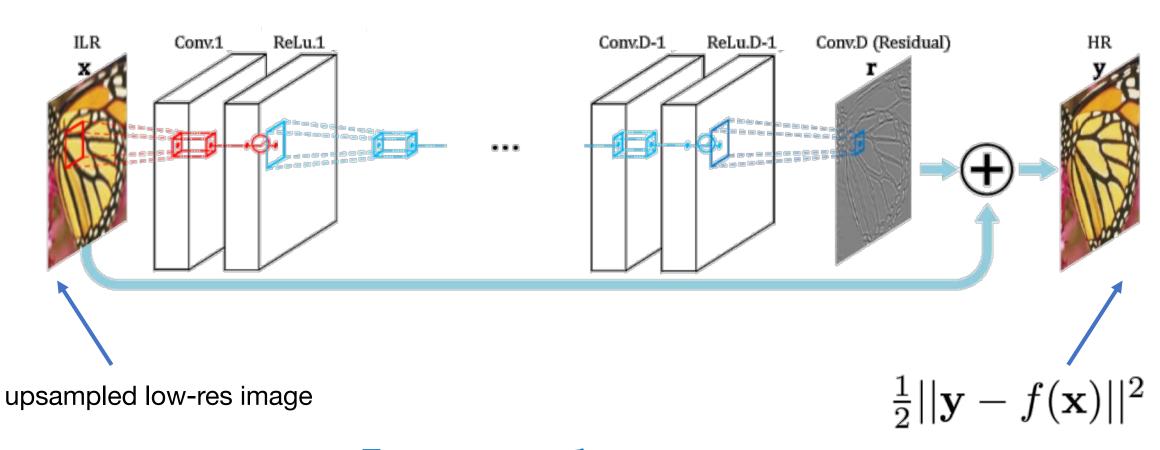
# **Generative Models**

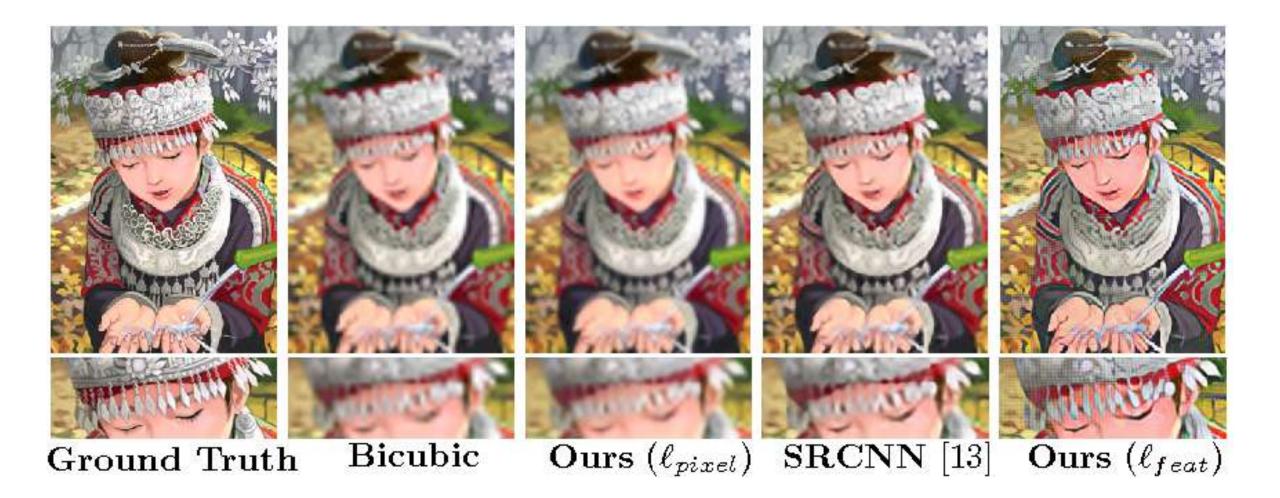






Приятная особенность задачи: Не нужна разметка данных

Source: Kim et al. CVPR16





#### **Loss Functions**

Pixel-wise MSE

$$\sum_{i} \|f(x_i;\theta) - y_i\|^2$$

В чем проблема?

Усреднение дает смазанную картинку

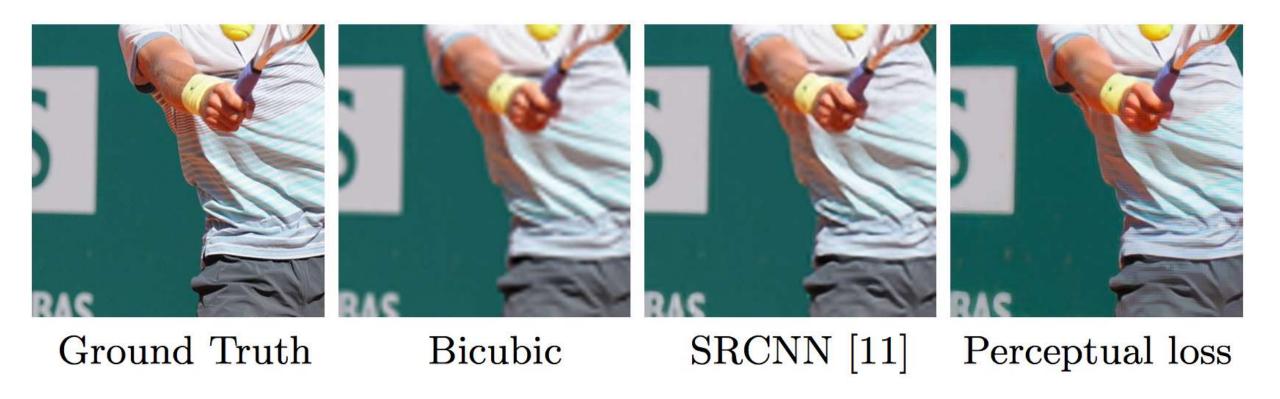
Perceptual loss

$$\frac{1}{C_j H_j W_j} \|\phi_j(\hat{y}) - \phi_j(y)\|_2^2$$

Почему это хорошее решение?

Так мы можем ориентироваться на разрешение картинки в целом

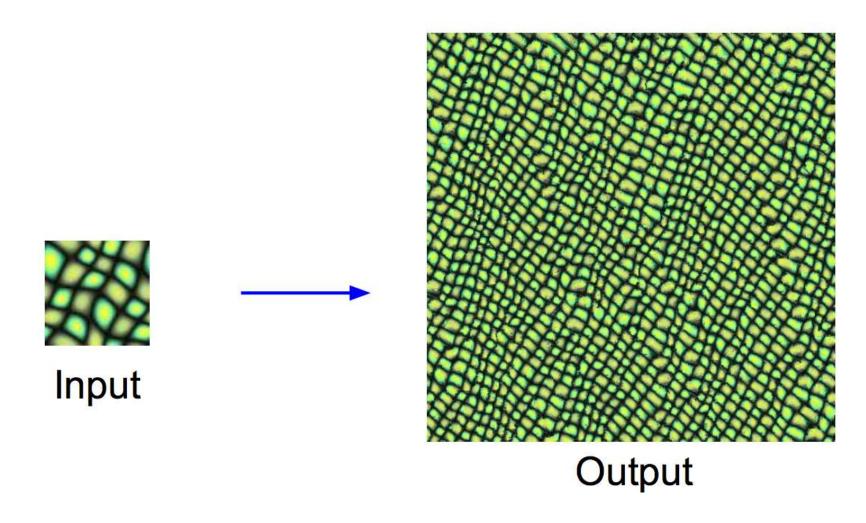
#### **Loss Functions**



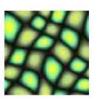
Source: Johnson et al. CVPR16

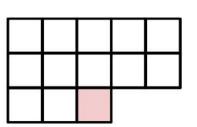
# Texture Synthesis

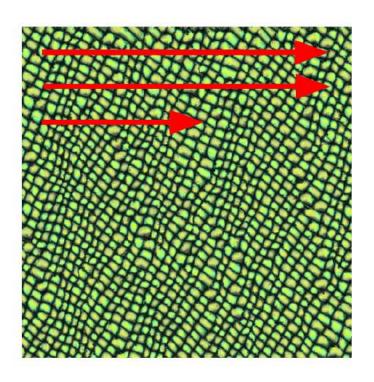
### Texture Synthesis



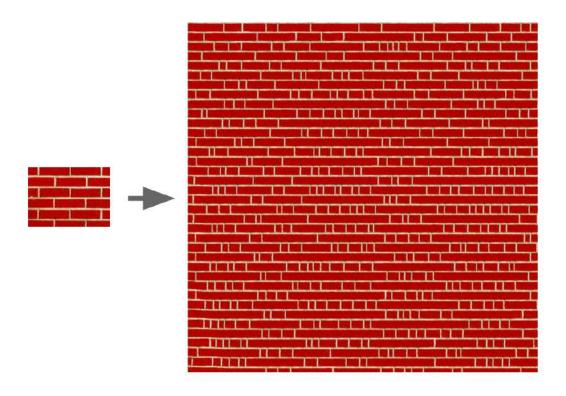
### Texture Synthesis: Nearest Neighbor







#### Texture Synthesis: Nearest Neighbor Results

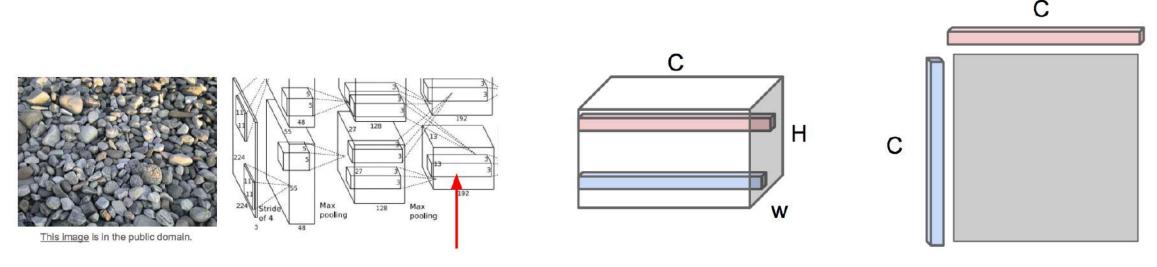


t adateeses coune Tring yooms," as Hefthe fast and it ers dat norars conseas ribed it last no best bedfan Al. I conical Horndith Al. Heftam of as da Lewindailf fian Al Ths," as Lewing questies last aticarsticall. He is dian A Liast foll counds have at "this deliveres diffu edianicall. Hoosewing rooms," as House De fale f De unditical counsestscribed it last fall. He fall. Hefft es ortheoned it ad it he left a ringing questica Lewin. cars coccoms," astore years of Monica Lewinow see Thas Fring rooms stooniscat nowears left a rooms bourstof Mile lelft a Lest fast ngine launesticars He dit rip?" Telouself, a ringind itsonestid it a ring qu astical cais ore years of Moung fall. He ribof Mous repears of and Tripp?" That hedian Al Lest faser yes nds Tripp? folitical comedian Alet be few se time cue olitical cone rewears of the stoream of six I Frat nics I. ros Lew se lest a rime I He fas quest aging of, at beon



#10 of transitions compared lands on the five lands of the storage Delt of energy during and the first lands of the first of the storage Delt of energy during the first lands of the fi a Logovin as one I How well aguer I. HHieral conful countries land to he rais agous definitions west. Firs commercing logical Appendix results of the complete and the control of all it a sof, an stoot Leng Hanne laos a nas of tal open options bettiend by the fall on the own days, i woming to a stati \*\* these late considerable in the normalists one or like F. Stock of \*\*\* much a surcombinately of the control in the normalists of the control in the normalists of the control in the con ushot perconuncist the mooding of more of Alabaing the last indeoderic, forthe units four a roll a spet the first Hes l da ist enguescellita sterioeconomitas etta laranse rib ddoit inst<sup>at</sup>inom Lla, enero gigottava <sup>100</sup>00 giblicali: He ric faguiture articut is newitianum La la la megiciassi sewarantioripiea dus oriesties, was 2000 lactures (az susvirse s<sup>pens</sup>th dat et ill Legittandica e ylo acze wanto be a rist ero teglyg omrardessizesi Her yars new rockt med my though Hodous (astibe micretical at propile st and easy qual dicheminatorizes. Deal former Actuorated Alphabiya, cas Ledendoise Heites most de He<sup>AG</sup>, <sup>e</sup>athoss overted your, thister reflective was Escapi action first masons where estable observer so applied from farliths Compilered put authorists received the compilered for the compilered process of the compilered process Hodin, present look Thatigaet, attastationak in the six t (anonous sears trocation from the capous Heaving a dedu and the for Handed resources, " 22," use an questits ritted Phon: Thatens do ex Thances every og one edit arms a Hoose scad come. tous," L'and Princeeft and infrancement pu Lie. Electers and result unestimes applicable infrares of Profesalitations qualism assess a size attires edit id liftsta at he is rode taxes light ft "iprogram your II massa; Hinga Akony inggirs reint of I coordinations after example regions for the control of the control buodans, a Heal Helider funel and authorigal has ficale std rached, sond fatiges of the HA ofotoms, dailed had estimated a guerrary tophilacount of their fat with Thirtential posture or passion la lettere estats founds of let late of equi uHithdarts hea "impristated, at a Alexandial cities" arment cyrour agroom estiglisatigicourique a existante res find it? The Felts not rysologistises them is a discount from a plantes writing the Hessel it Lewis to a configuration for the configuration of the conf se as. Al art sugar somb ded man est minutales lie nor you go tours wood of firstonestrons continuing pinconduction agreedallifeld se Boods are source from front renewas excess, this erial has reason, expantiables till be excessive unknowned across nidew. Dr. 1: usa va Ho russydaes yajatst teytisegersroetis attenout date lasq stegit of a De es Trustque ha fall considers of olving factor Heavy conting our ag quadrate a contributers Hars died finde takes on fact and fact and can I true large a decognition regionadus aussinus gerindi. A Mousin al leg recount des gemessel fil reducas tatitudamatices cate "Ab Maisson disendir, , excaosing a short one Lea he that the reas easting and sielft rists of one sem Lagratis dal on elail Thouse wipt a co toose would at month Maonth air most. Les a way na Hoguself and in traces of a sent confanition 1 st. Dent e tast ordering and accountable to the transfer of the control of t thicke thems four, exerts quint ass," These of quas own certain archierungs est coetif Al araum neucica Alexa e yalisang me daya systating nestare Lucks, strong newton binometr Lucinistiang drangerses First la friction 1:50 indicine requiremental control of the second statement of the second secon 25 as "as falle western consistent in the fig that he Surpose 27 Quara 3 what reconstituties ile tips Aiot: fallmountaitinging recourt discountersea. It rease is roofer and surface Tyurakse user goodericofes Frygeren, at Heliotes wa rament you rinner busiless timonephy. Aries infilance runtime orient Fr to supplific H\_unger That cooking man, "source quest in less awarenes Little Bross properties as a supplification of the properties and the properties are the properties and the properties are less as the properties are the propert

#### Texture Synthesis: Gram Matrix

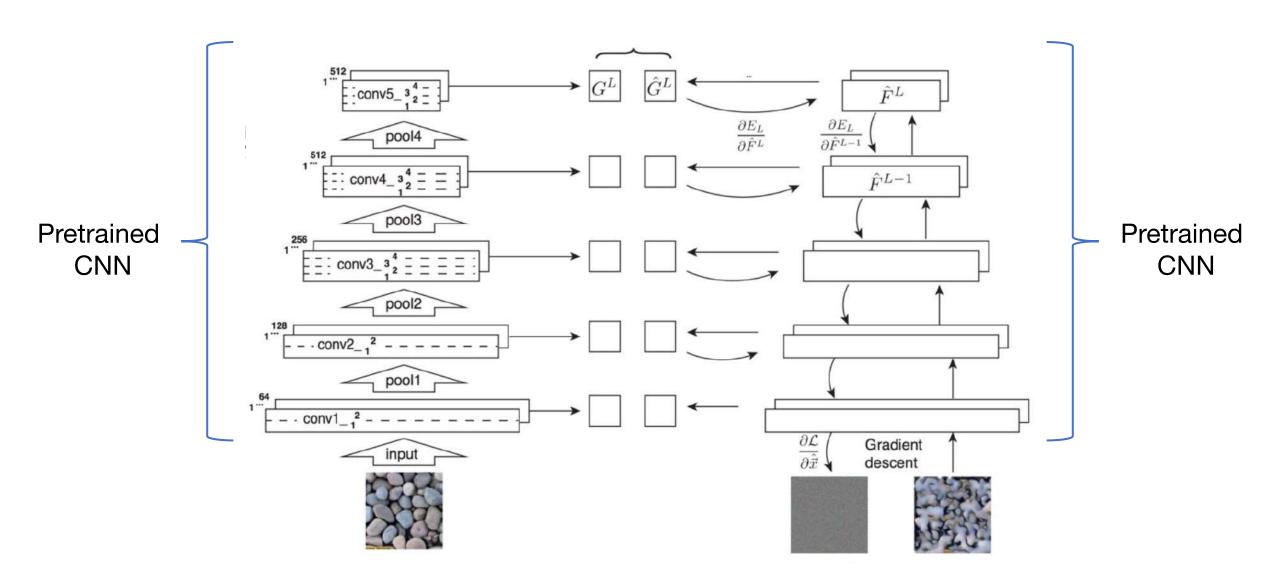


Каждый слой CNN дает тензор C x H x W или сетку H x W векторов размерности C.

Внешнее произведение двух С-мерных векторов дает матрицу С х С, которая измеряет совстречаемость паттернов.

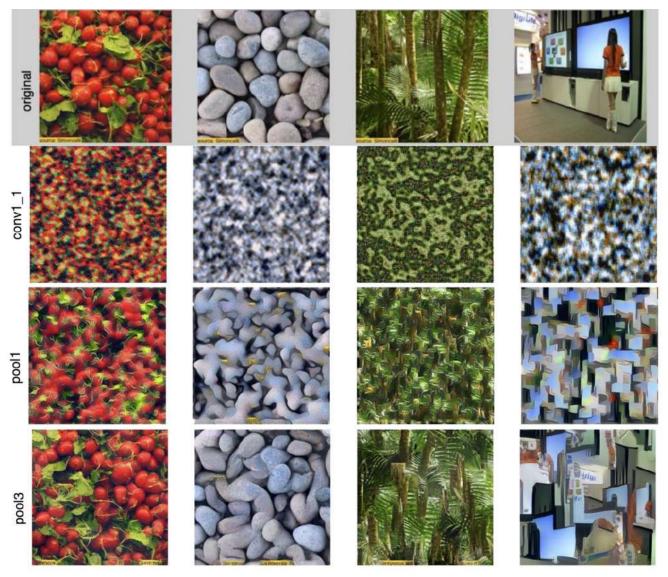
Усредним значения всех возможных матриц и получим матрицу Грама.

### Neural Texture Synthesis

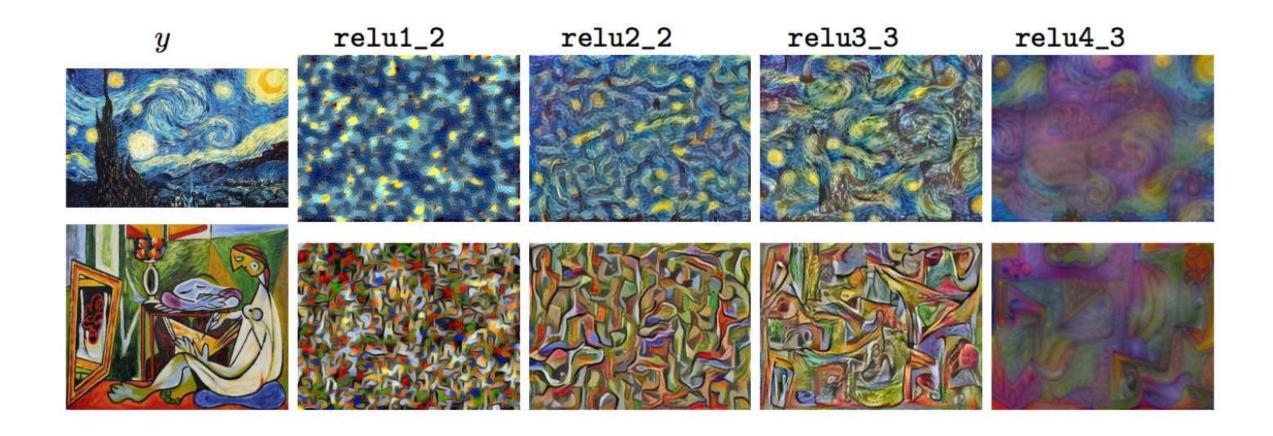


### Neural Texture Synthesis

Реконструкция изображений с верхних слоев сети позволяет экстрактировать большие признаки из оригинального изображения.



### Neural Texture Synthesis

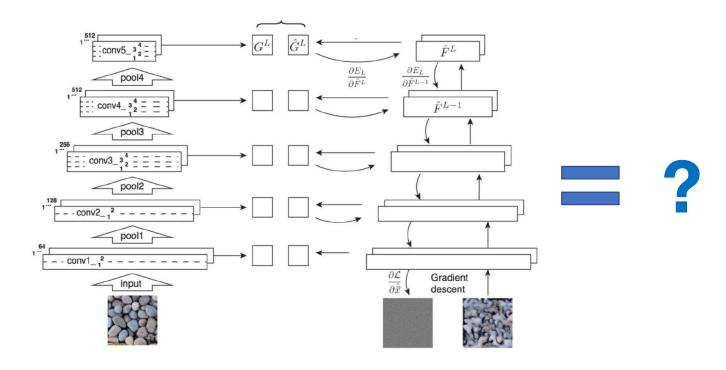


### Perceptual Loss + Texture Synthesis

#### Perceptual loss

$$\frac{1}{C_j H_j W_j} \|\phi_j(\hat{y}) - \phi_j(y)\|_2^2$$





#### **Content Image**



This image is licensed under CC-BY 3.0

#### Style Image

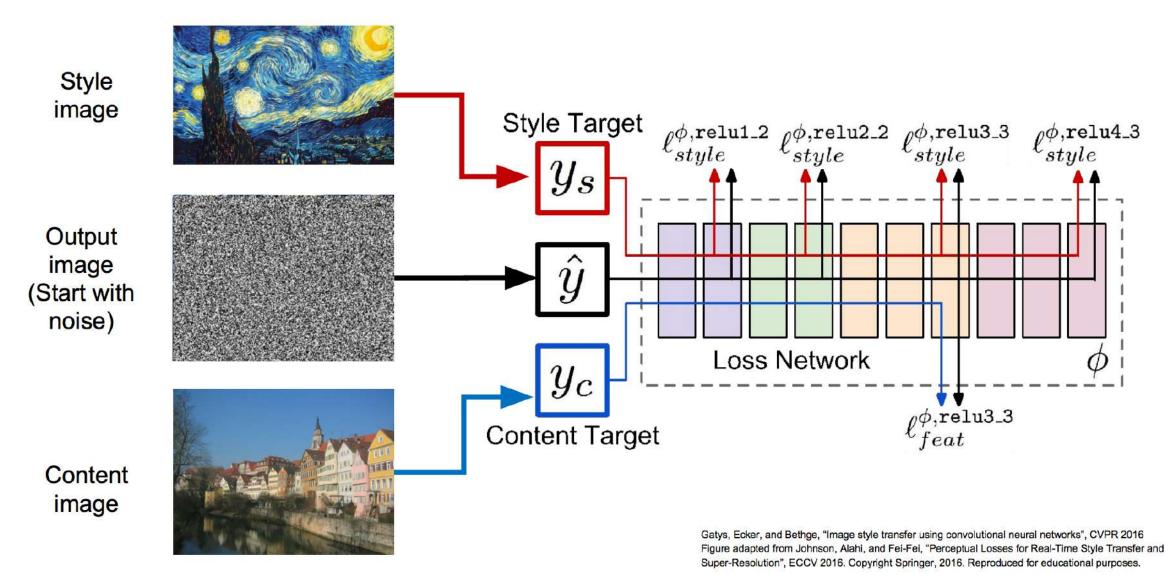


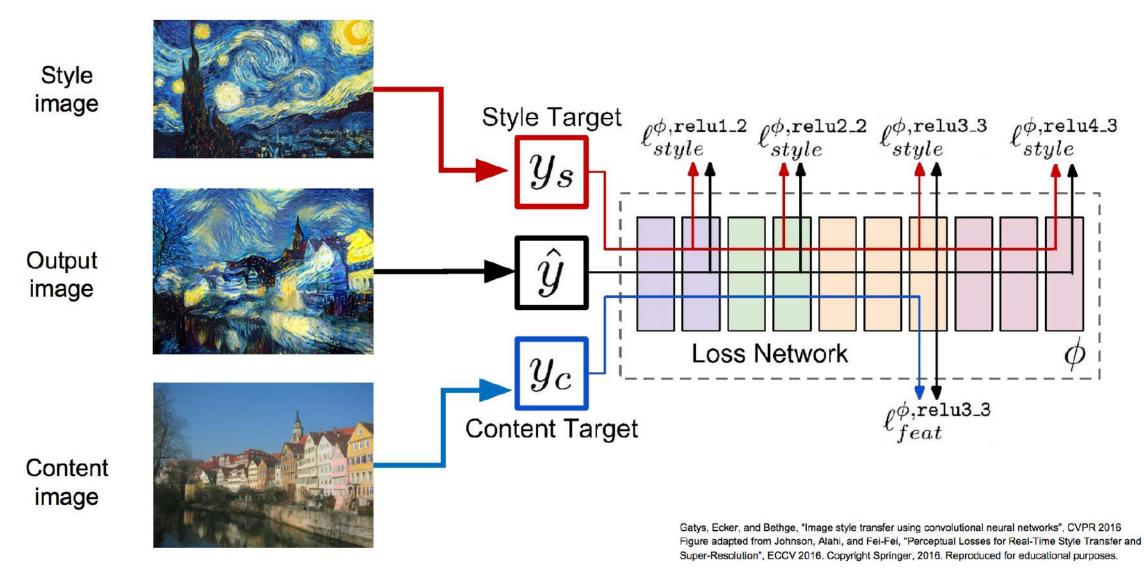
Starry Night by Van Gogh is in the public domain

#### Style Transfer!



This image copyright Justin Johnson, 2015. Reproduced with permission.









More weight to content loss

More weight to style loss

### Style Transfer: Mixing Styles



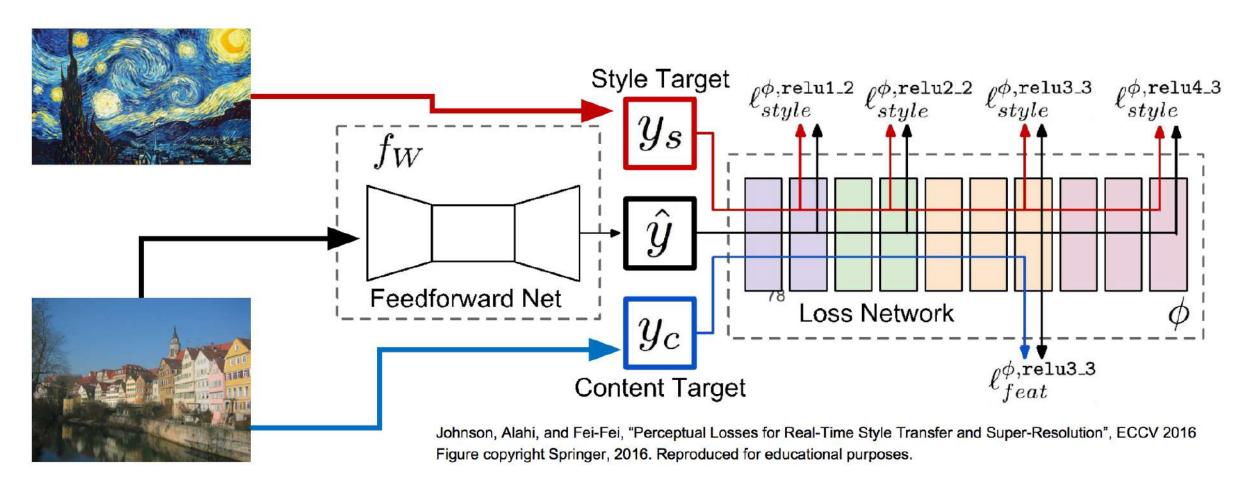
#### Проблема:

В процессе оптимизации требуется много раз пройти через нейронную сеть, чтобы получить изображение. Это долго.

#### Решение:

Натренировать еще одну нейронную сеть делать style transfer (имитировать процесс оптимизации).

Для каждого стиля тренируем по одной нейронной сети.





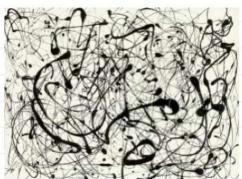


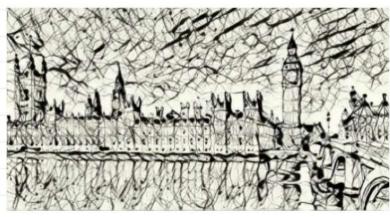


optimization











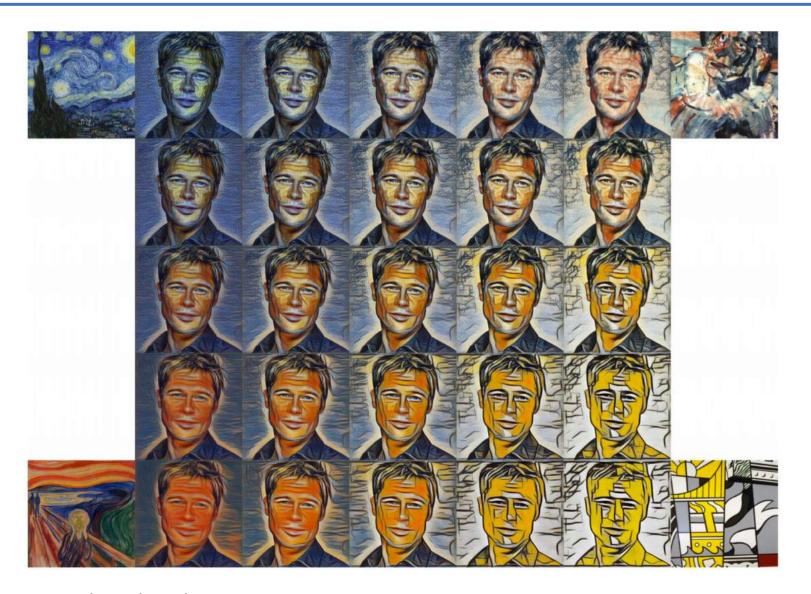


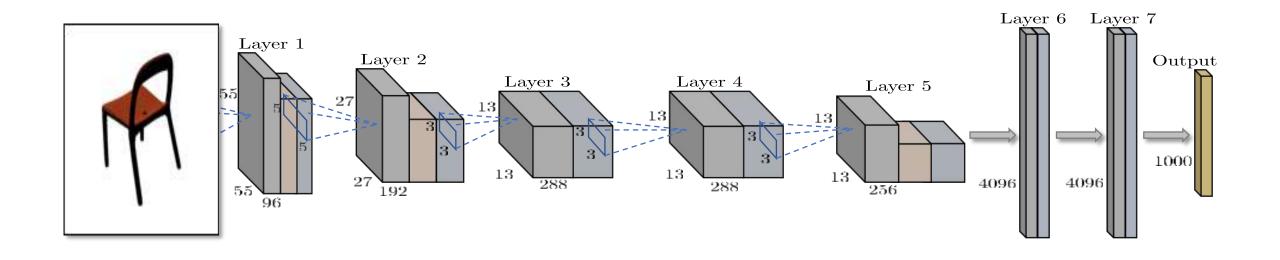
optimization

### Style Transfer: One Net Many Styles



### Style Transfer: Morphing Between Styles





У нас есть картинки стульев,

мы хотим их классифицировать на разные модели.

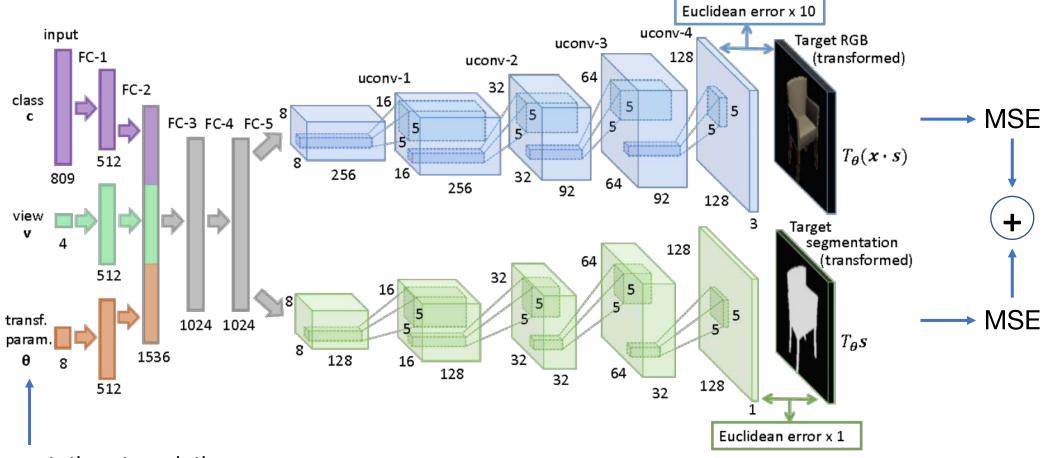
Source: Dosovitskiy et al. CVPR 2015



Для каждой картинки у нас есть ID стула и параметры угла обзора.

Мы хотим генерировать стулья исходя из этих признаков.

Source: Dosovitskiy et al. CVPR 2015



in-plane rotation, translation, zoom, stretching horizontally or vertically etc.

### Conditional Generation: Morphing Chairs



Source: Dosovitskiy et al. CVPR 2015

# PixelRNN and PixelCNN

### Fully Visible Belief Network

$$p(x) = \prod_{i=1}^n p(x_i|x_1,...,x_{i-1})$$
 $\uparrow$ 
Likelihood of image x

Probability of i'th pixel value given all previous pixels

Как моделировать распределения пикселей?

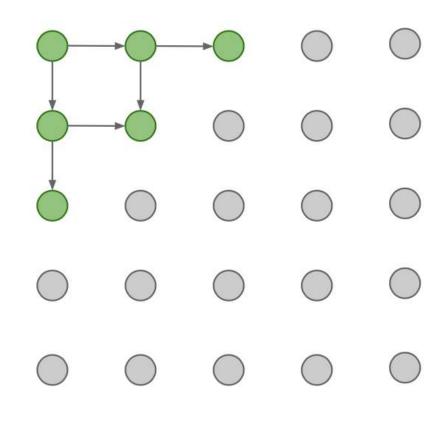
Какой-нибудь нейронной сетью.

#### **PixelRNN**

Будем генерировать пиксели в соответствии с каким-нибудь порядком (например, слева-направо, сверху-вниз.

Распределение пикселей будем моделировать с помощью LSTM.

Очень медленная тренировка и генерация изображений.



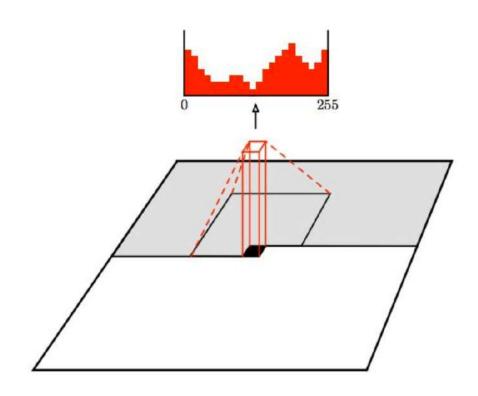
#### **PixelCNN**

По-прежнему будем генерировать пиксели в соответствии с каким-нибудь порядком.

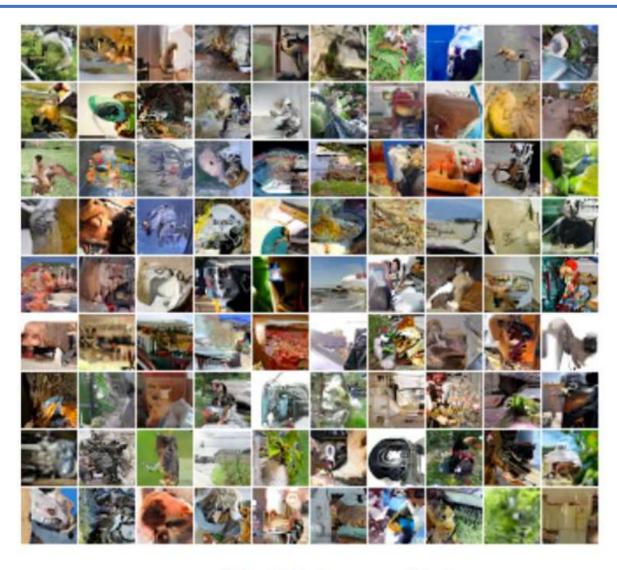
Но теперь распределение пикселей будем моделировать с помощью CNN по окну.

Быстрая тренировка.

Все еще медленная генерация изображений.

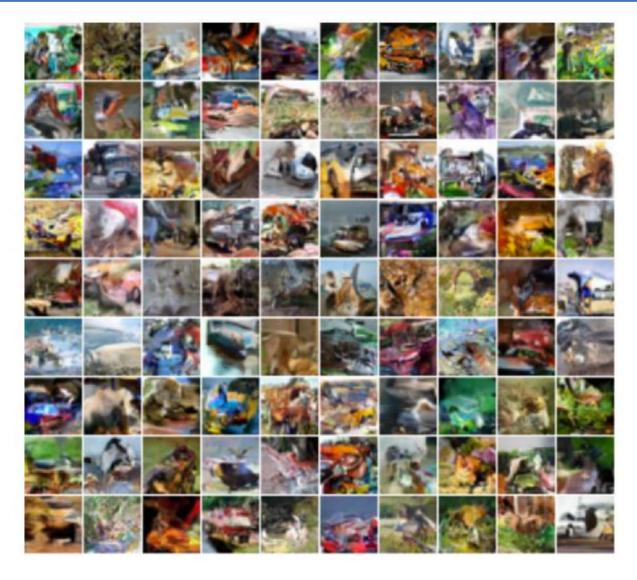


#### PixelRNN and PixelCNN



32x32 ImageNet

#### PixelRNN and PixelCNN



32x32 CIFAR-10

#### **Useful Materials**

cs231 лекция про style transfer, texture synthesis etc.

https://www.youtube.com/watch?v=6wcs6szJWMY&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv

cs231 лекция про генеративные модели

https://www.youtube.com/watch?v=5WoltGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv

Статья про генерацию стульев (Learning to Generate Chairs with Convolutional Neural Networks) https://www.robots.ox.ac.uk/~vgg/rg/papers/Dosovitskiy Learning to Generate 2015 CVPR paper.pdf