

Глубокое обучение и вообще

Ульянкин Филипп

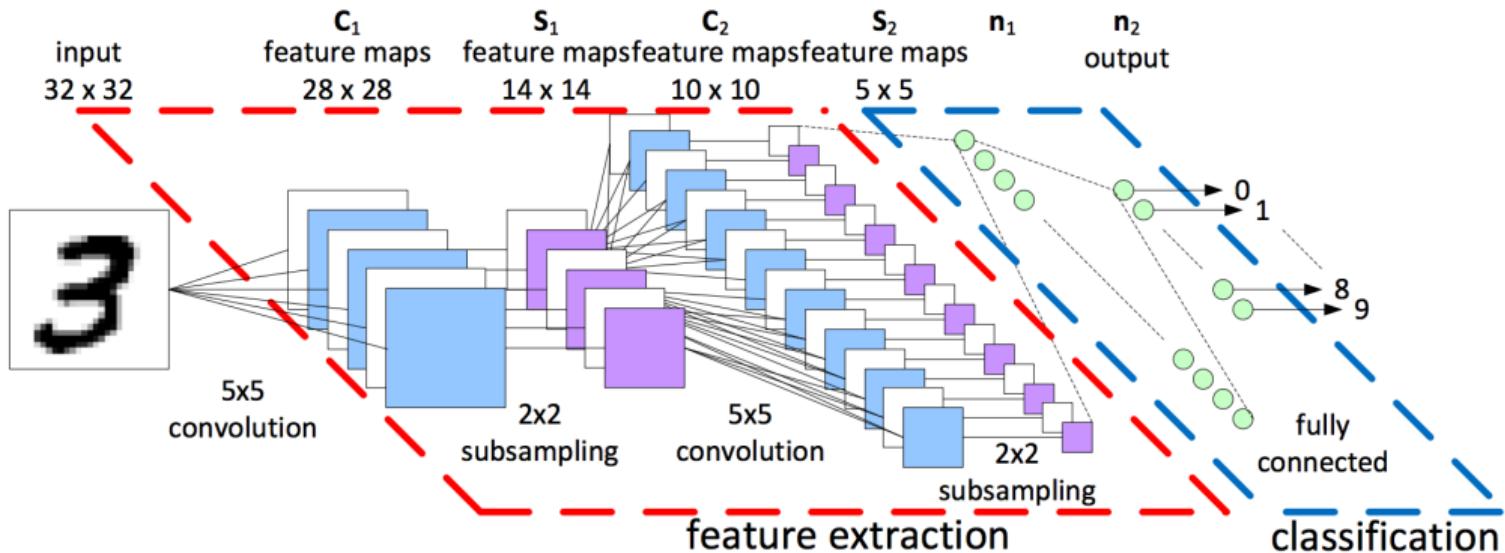
15 июня 2022 г.

Посиделка 10 (мб 11): Задачи компьютерного зрения

Agenda

- Локализация и сегментация
- Перенос стиля

Классификация изображений

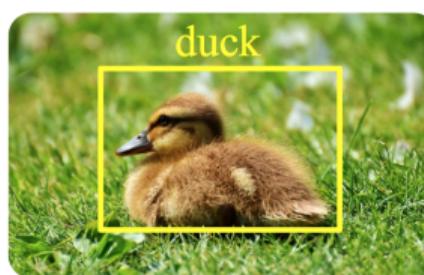
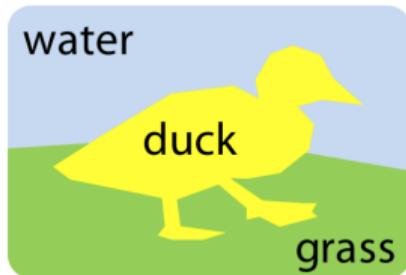


Обычно хочется другого

Semantic segmentation:



Object classification
+ localization:

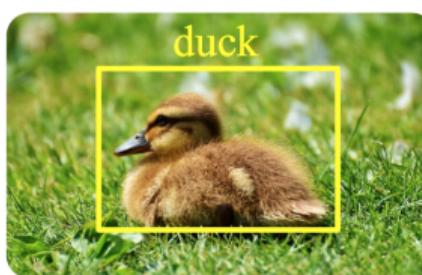
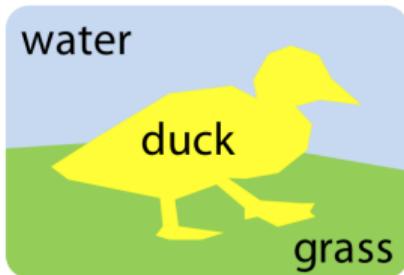


Обычно хочется другого

Semantic segmentation:



Object classification
+ localization:



Мы будем сводить эти задачи к классификации

Сегментация

Постановка задачи

Image



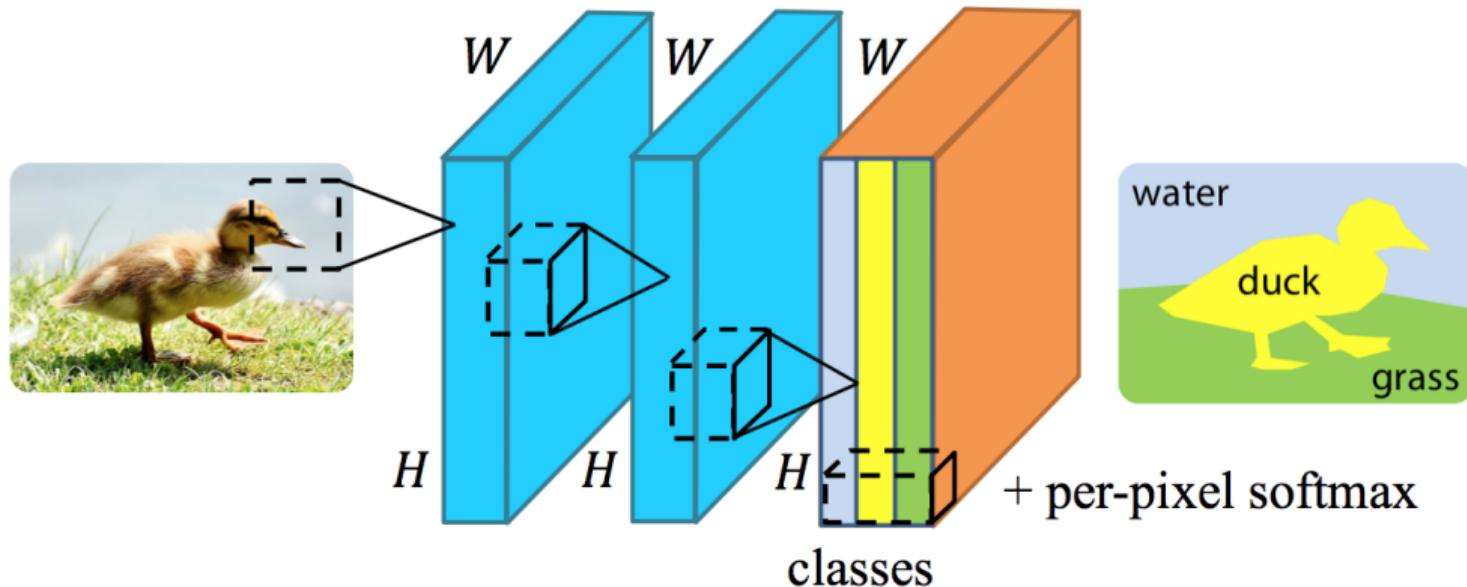
Objects



- На вход модели подаются изображения и их корректные сегментации
- Каждый объект содержит сильно больше информации, чем при классификации
- Можно обучаться на более маленьких выборках

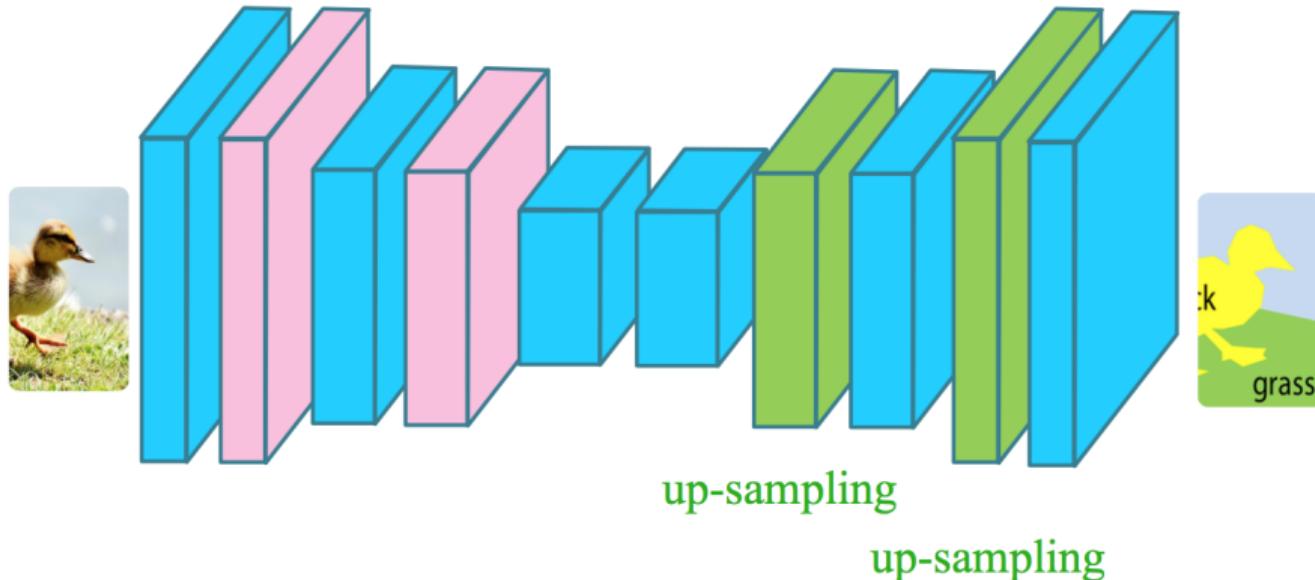
<http://host.robots.ox.ac.uk/pascal/VOC/voc2012/>

Сегментация



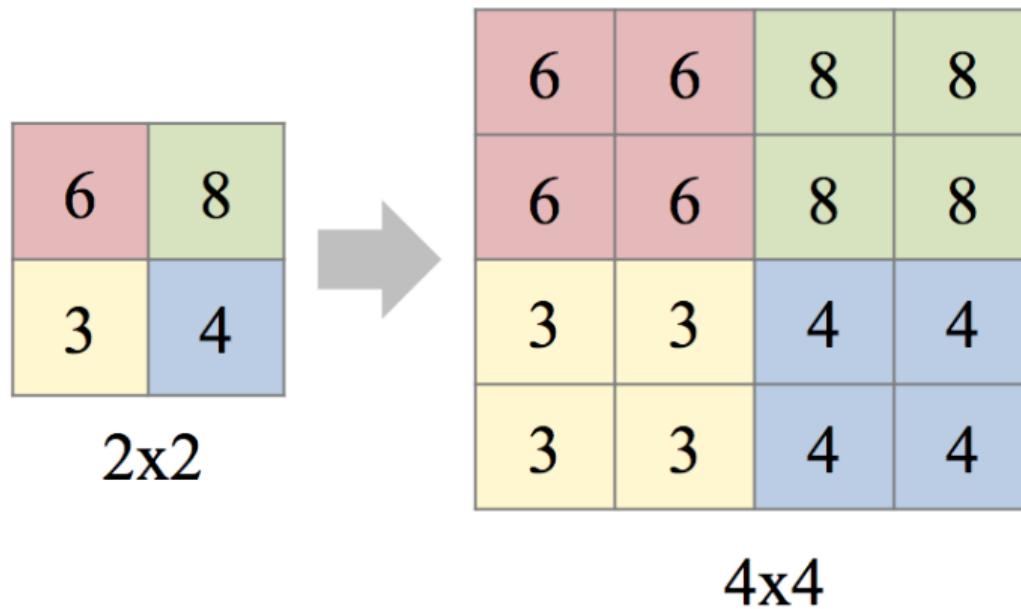
- Нам нужно научиться классифицировать каждый пиксель
- Куча свёрток и попиксельный softmax без пулинга (наивный подход)

Сегментация

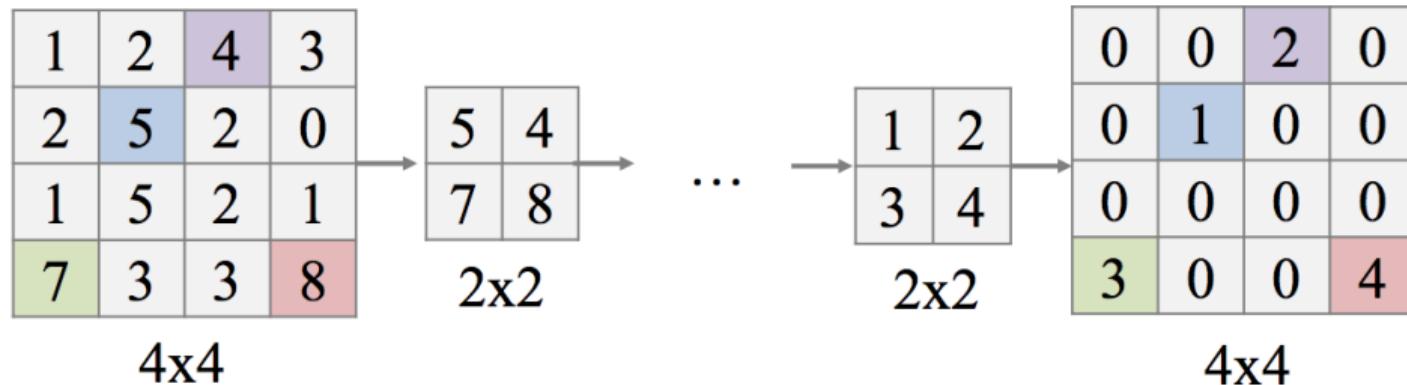


- Если захотим добавить пулинг, придётся делать анпулинг

Nearest neighbor unpooling

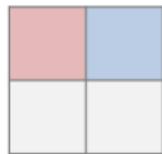


Max unpooling

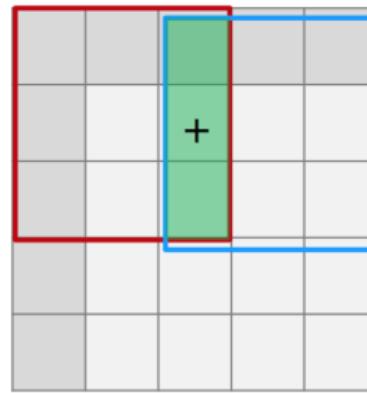


Learnable unpooling: Transpose convolution

Input: 2x2



Input gives
weight for
filter

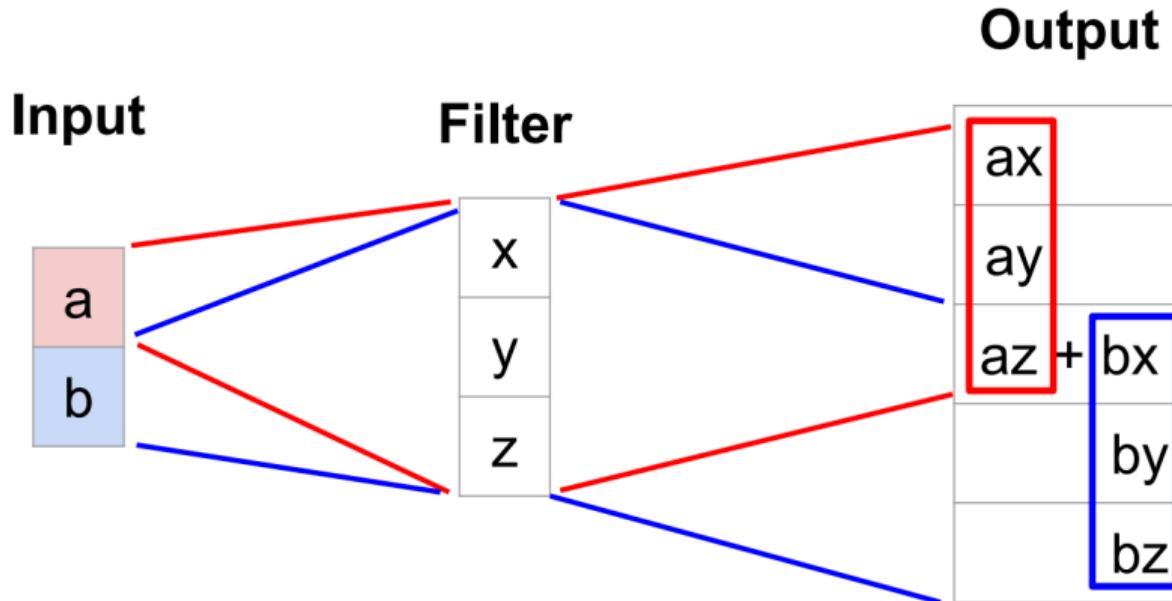


Stride: 2

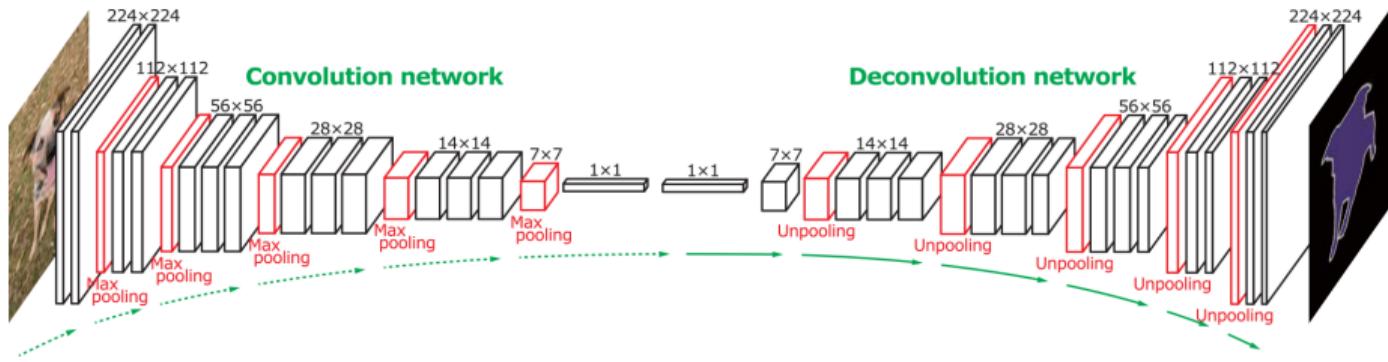
Output: 4x4

- Каждую клетку надо распаковать в 4 клетки \Rightarrow свёртка 3×3 со сдвигом 2

Пример:

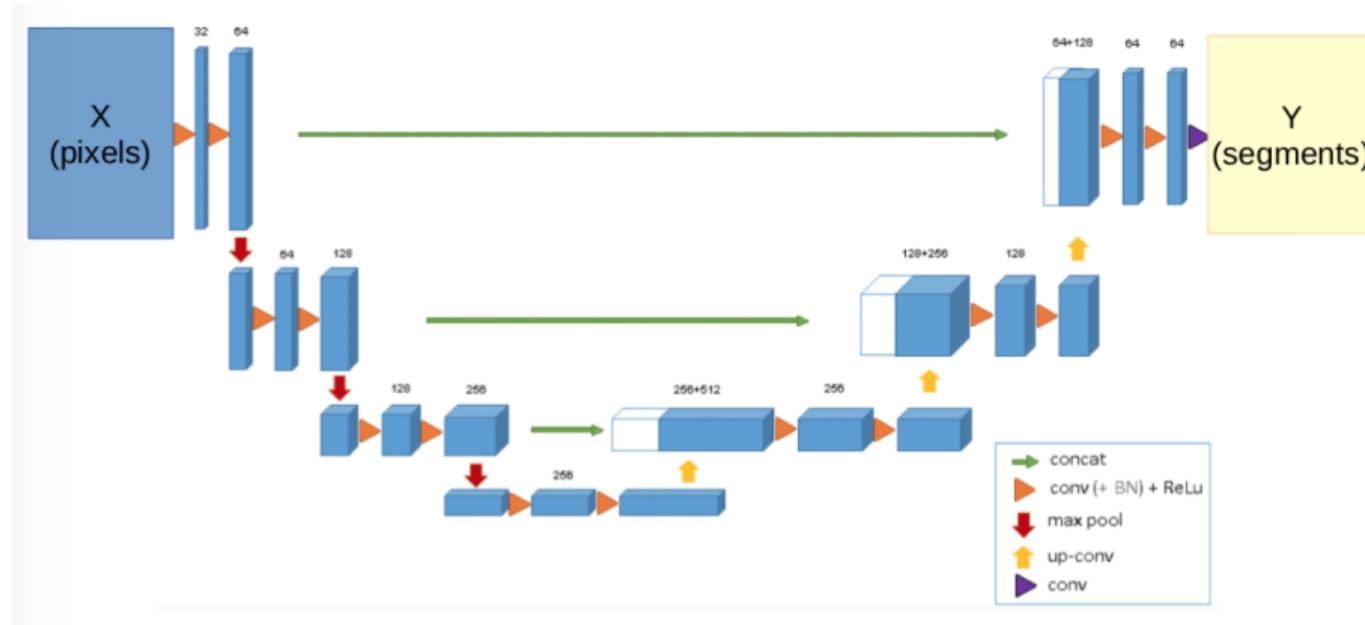


Fully-convolution net



- Свернули в скрытое представление, развернули, спрогнозировали

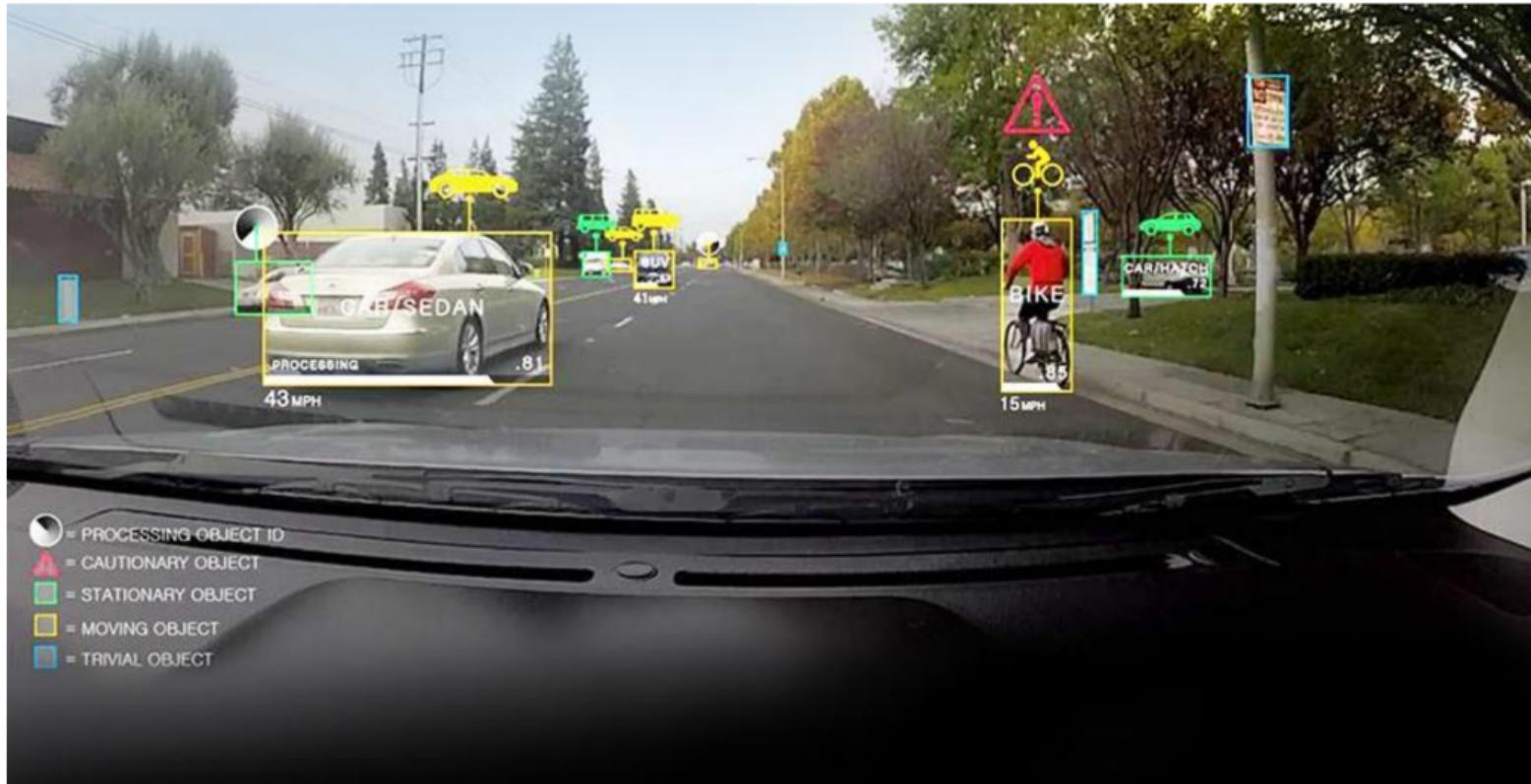
U-net



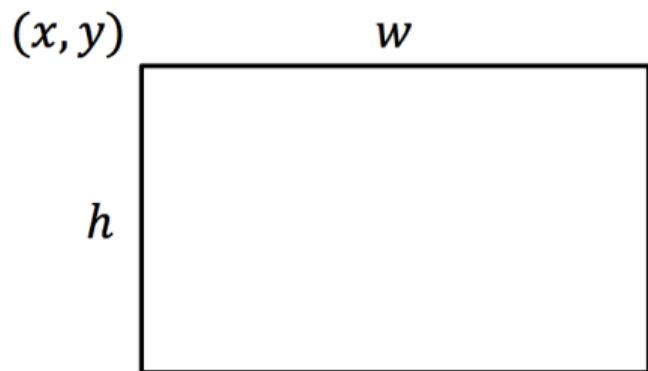
- Можно добавить связи между слоями, отражающими одинаковую абстракцию, это должно улучшить модель

Детекция объектов (локализация)

Детекция объектов (локализация)

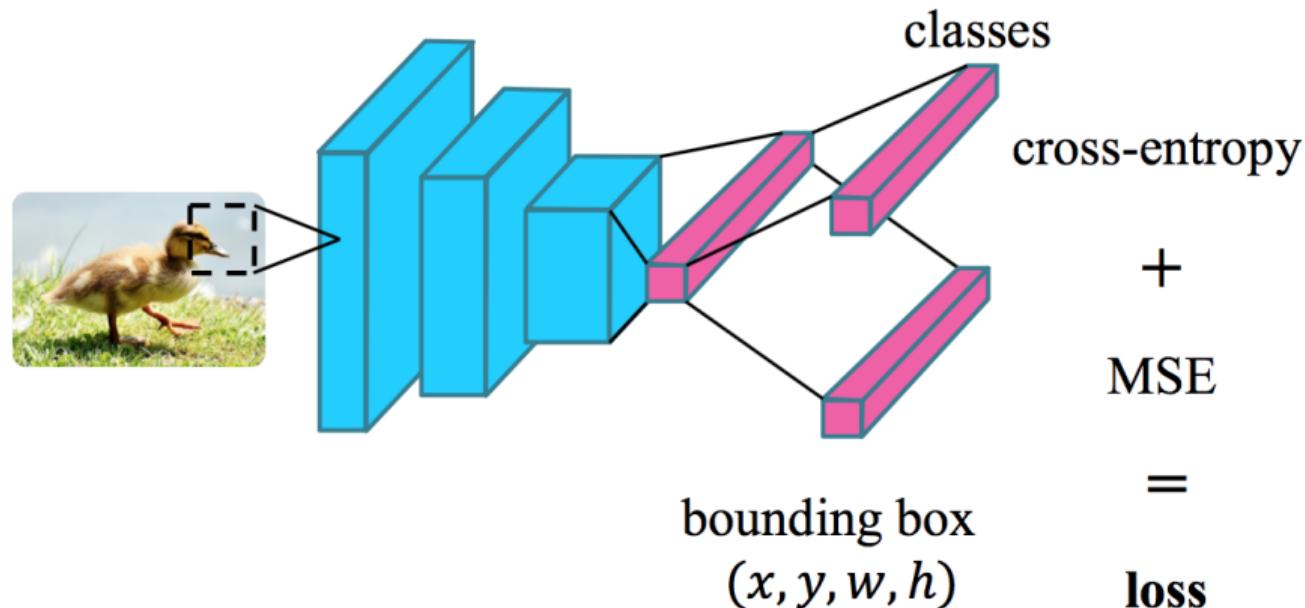


Детекция объектов (локализация)



- для локализации объекта нужно нащупать рамочку, в котором он находится
- рамочка описывается параметрами (x, y, w, h)

Детекция объектов (локализация)



<https://arxiv.org/abs/1506.02640>

Детекция объектов (локализация)

- Перебирать все прямоугольники очень долго
- Не очень понятно, что такое хороший прямоугольник
- Прямоугольники могут быть разного размера
- Можно сначала генерировать кандидатов в прямоугольники с помощью классических методов компьютерного зрения, а потом выбирать среди них лучшего

<https://arxiv.org/abs/1311.2524>

Детекция объектов (локализация)

- Перебирать все прямоугольники очень долго
- Не очень понятно, что такое хороший прямоугольник
- Прямоугольники могут быть разного размера
- Одновременный поиск кандидатов и определение их классов (YOLO, you only look once)

<https://lilianweng.github.io/lil-log/2018/12/27/object-detection-part-4.html>

Перенос стиля

Приложение Prisma



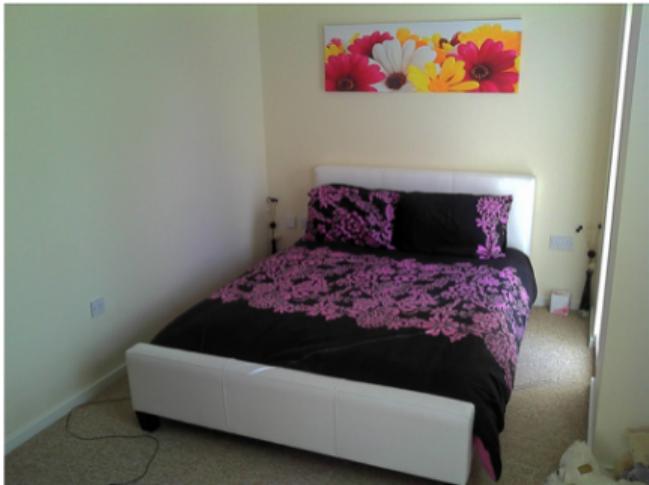
Перенос стиля



Перенос стиля



Ваша комната



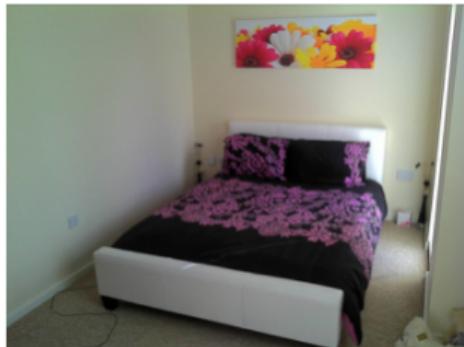
Комната сына
маминой подруги



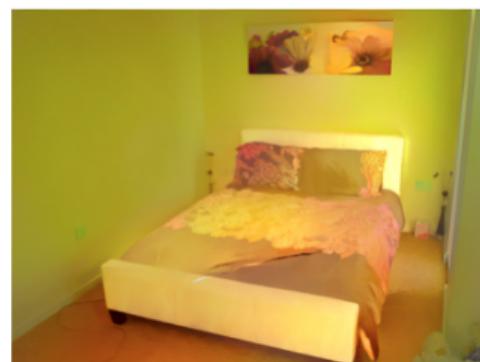
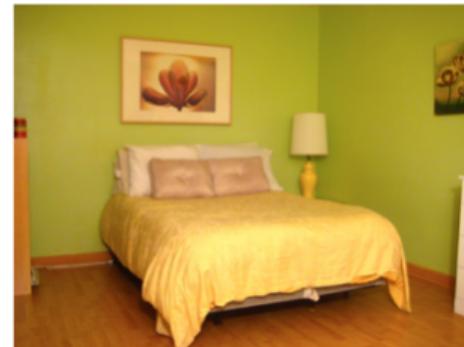
<https://habr.com/ru/post/402665/>

<https://github.com/LouieYang/deep-photo-styletransfer-tf>

Ваша комната



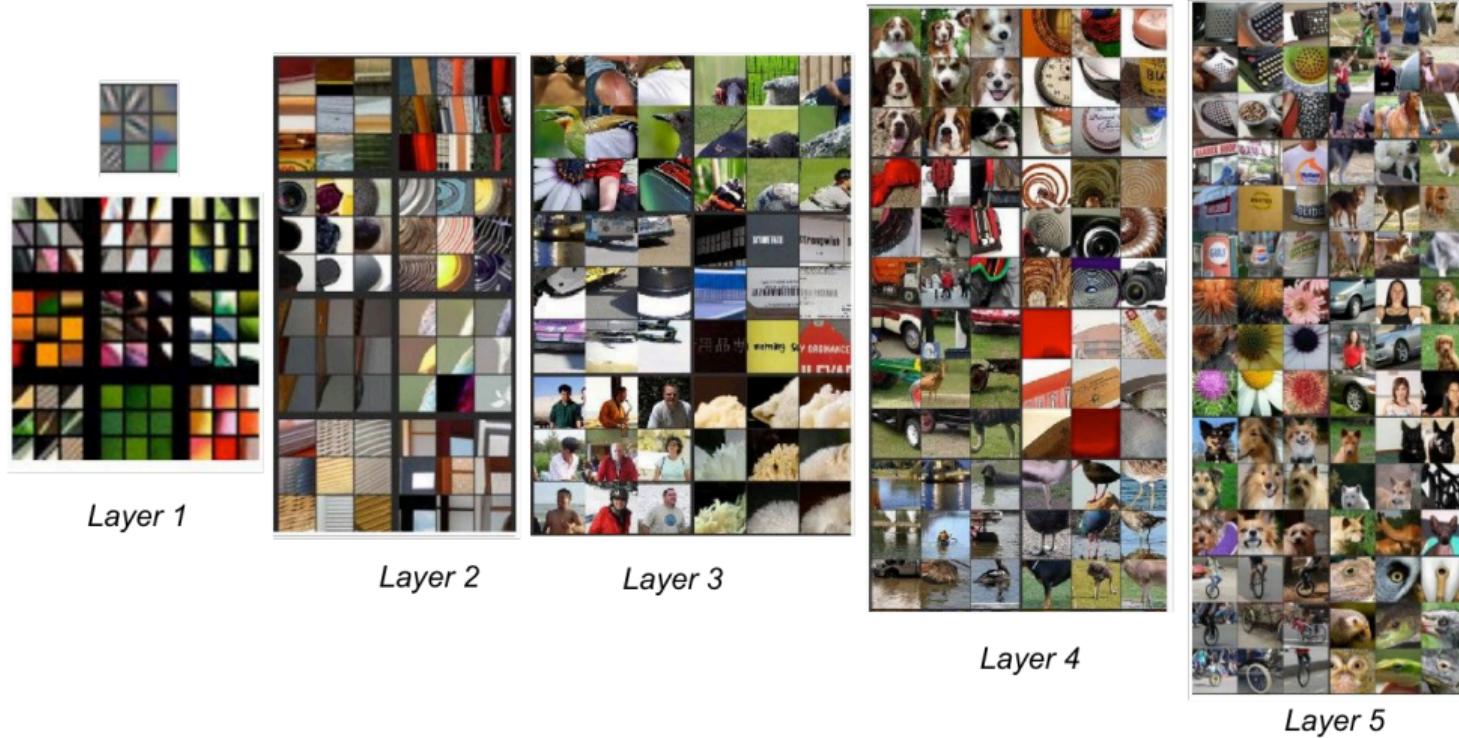
Комната сына
маминой подруги



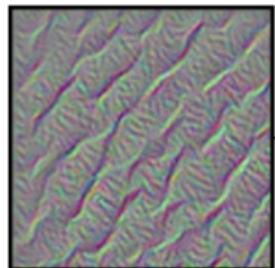




Опять этот слайд!



Есть несколько способов заглянуть внутрь свёрточной сетки, мы сейчас посмотрим на один из них

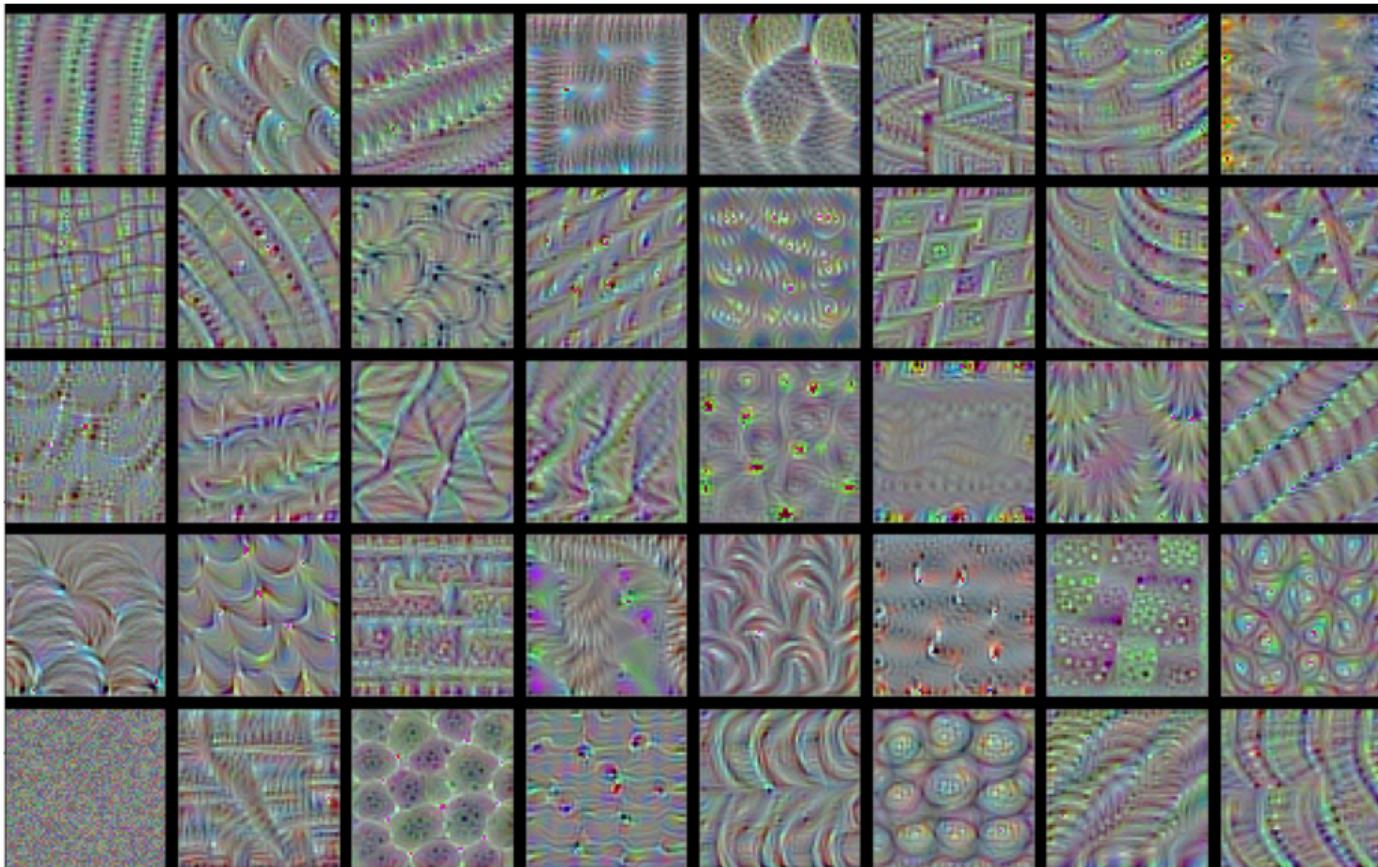


Inception v3 Model

- Orange square: Convolution
- Blue square: AvgPool
- Green square: MaxPool
- Pink square: Concat
- Dark blue square: Dropout
- Purple square: Fully connected
- Red square: Softmax

Use gradient for a feature to update input image

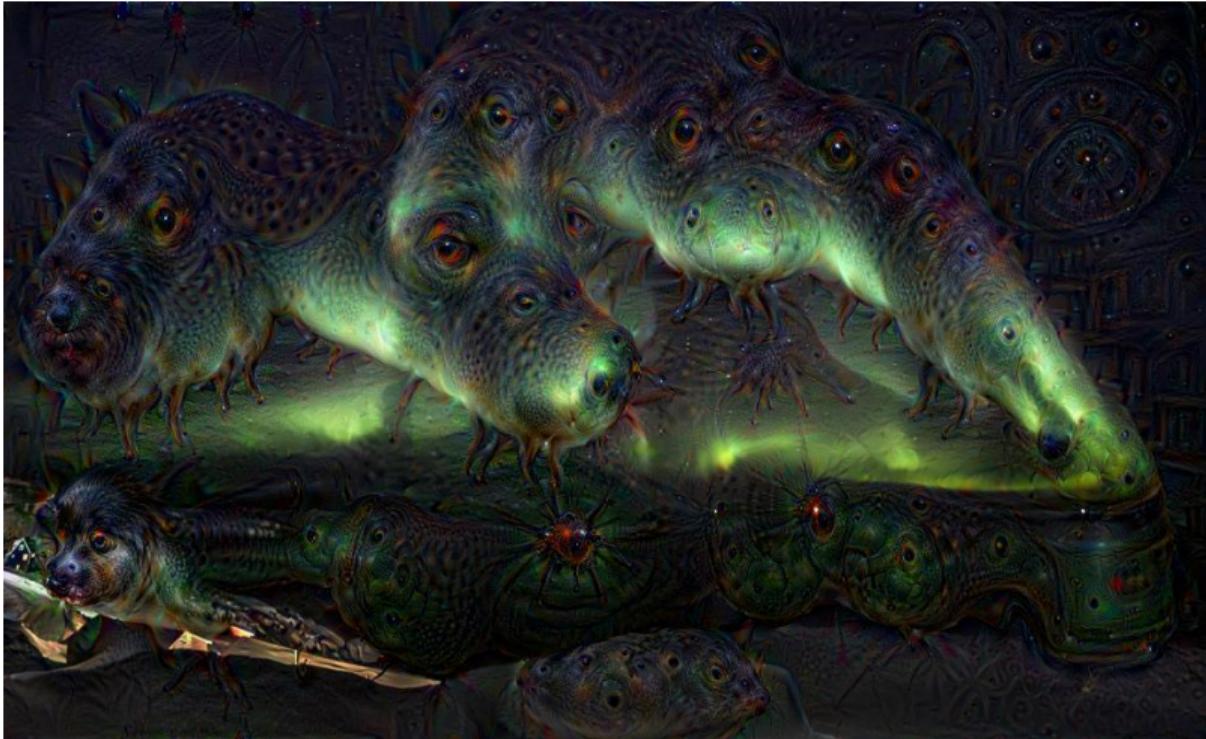
Пример из вашей домашки



Что видит свёртка

- На вход идет пустое изображение, мы хотим изменить его пиксели так, чтобы активация конкретной свёртки была максимальной
- Максимизируем среднее значение свёртки по пикселям
- Шаг градиентного спуска: меняем пиксели так, чтобы свёртка выдавала на выход более большие значение
- На входной карточке постепенно прорисовывается шаблон, который возбуждает соответствующую свёртку
- Если на вход в сетку подсунуть не пустую карточку, а какое-то изображение, то фильтр отрисуется на нём. Если эту процедуру немного подправить, получится наркомания под названием **Deep dream**

Deep dream



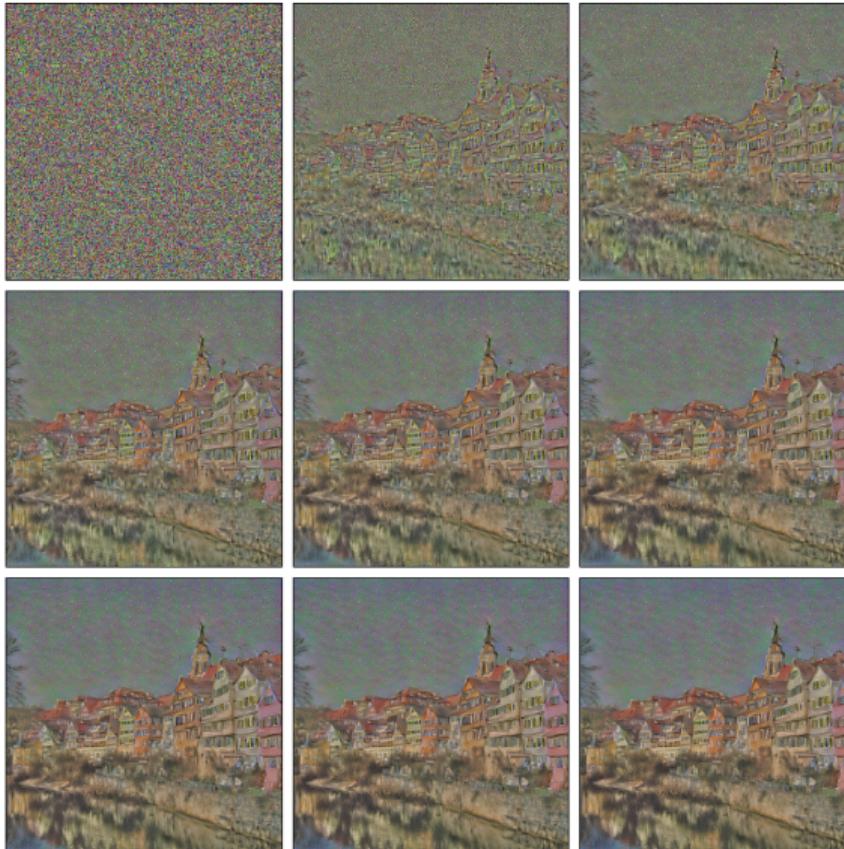
<https://nplus1.ru/material/2015/07/13/use>

Content loss



<https://habr.com/ru/company/mailru/blog/306916/>

Content loss

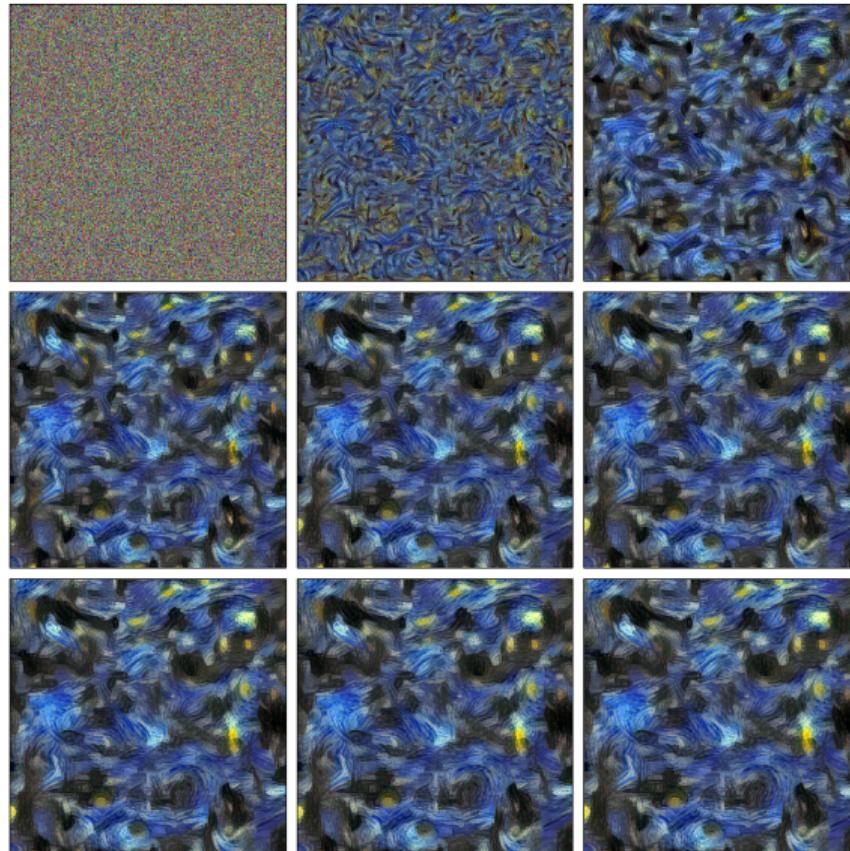


Style loss

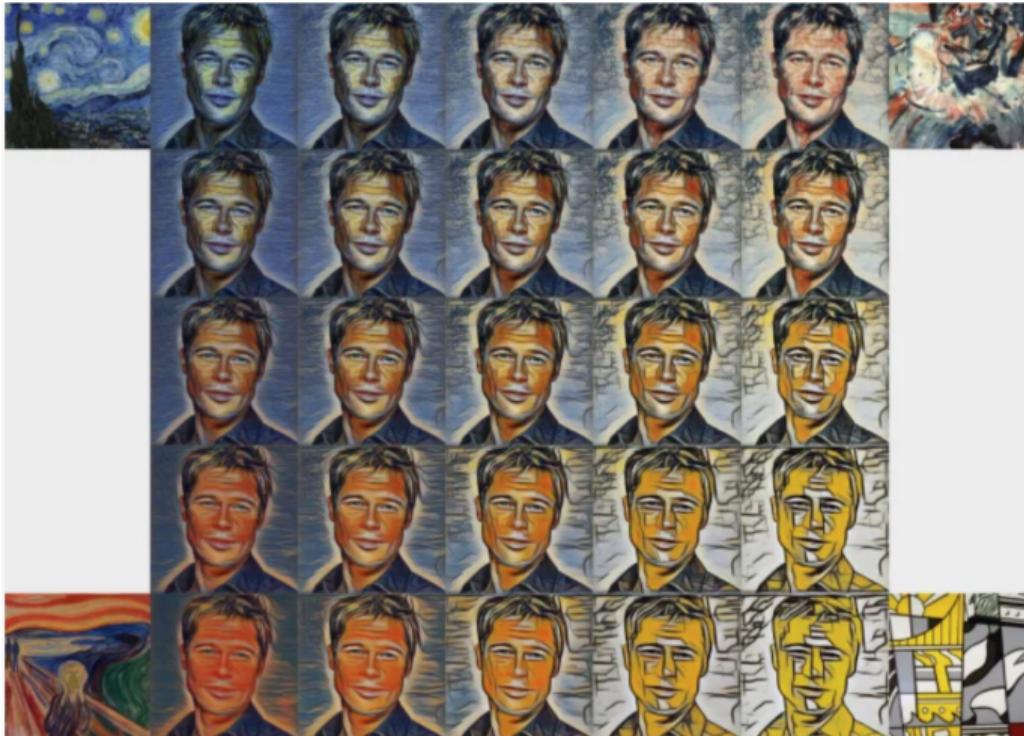


<https://habr.com/ru/company/mailru/blog/306916/>

Style loss



Смесь нескольких стилей



<https://arxiv.org/pdf/1610.07629.pdf>

Переносим стиль!