

Глубокое обучение и вообще

Никита Бекезин

27 октября 2021 г.

Неделя 8: Что видят сетки, перенос стиля, локализация, сегментация

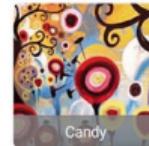
Agenda

- Что видят свёрточные сетки?
- Перенос стиля
- Локализация и сегментация

Что видят нейросетки

Перенос стиля

Приложение Prisma



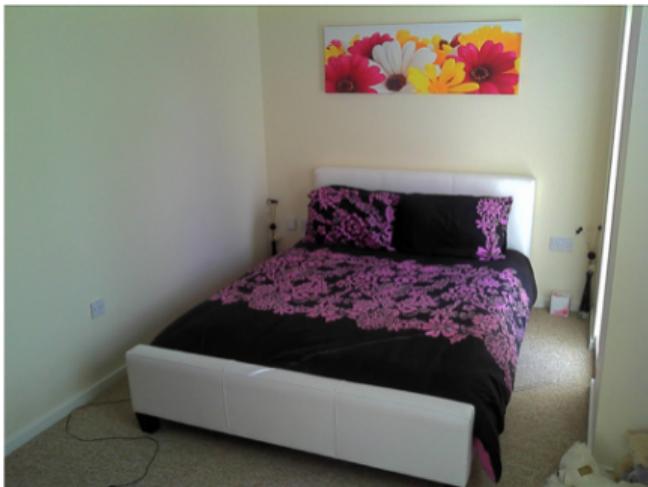
Перенос стиля



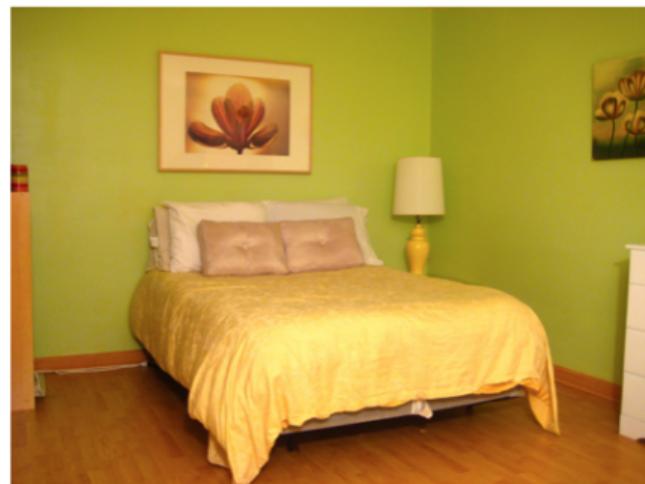
Перенос стиля



Ваша комната



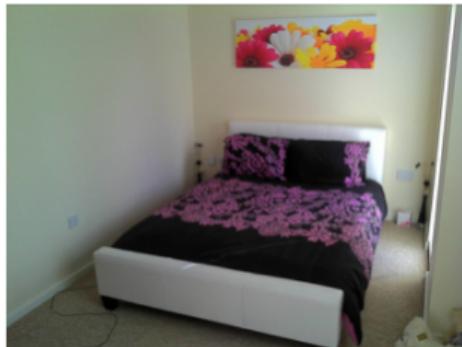
Комната сына
маминой подруги



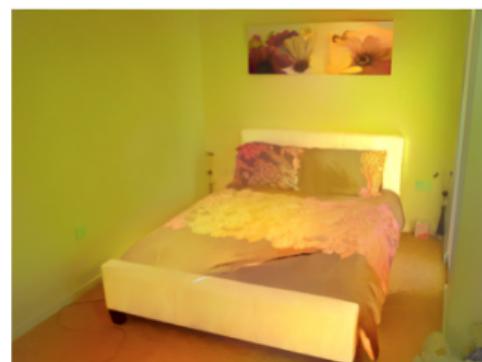
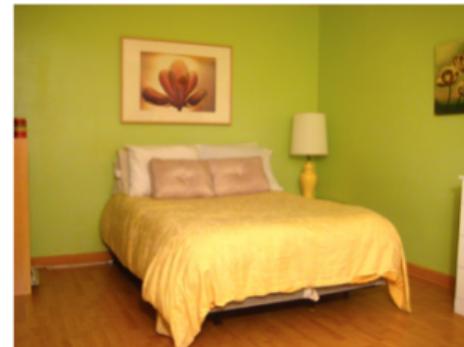
<https://habr.com/ru/post/402665/>

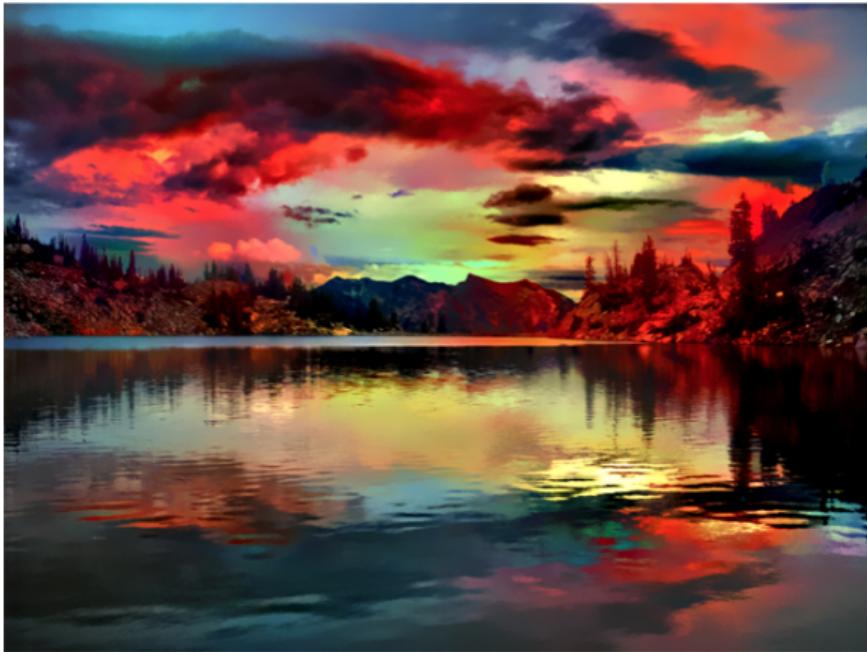
<https://github.com/LouieYang/deep-photo-styletransfer-tf>

Ваша комната



Комната сына
маминой подруги







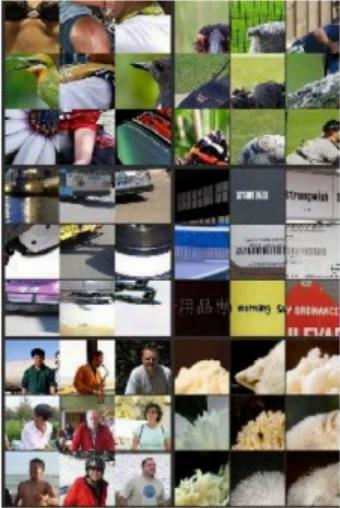
Опять этот слайд!



Layer 1



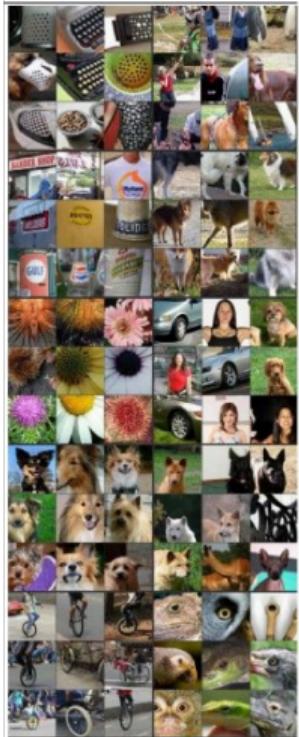
Layer 2



Layer 3

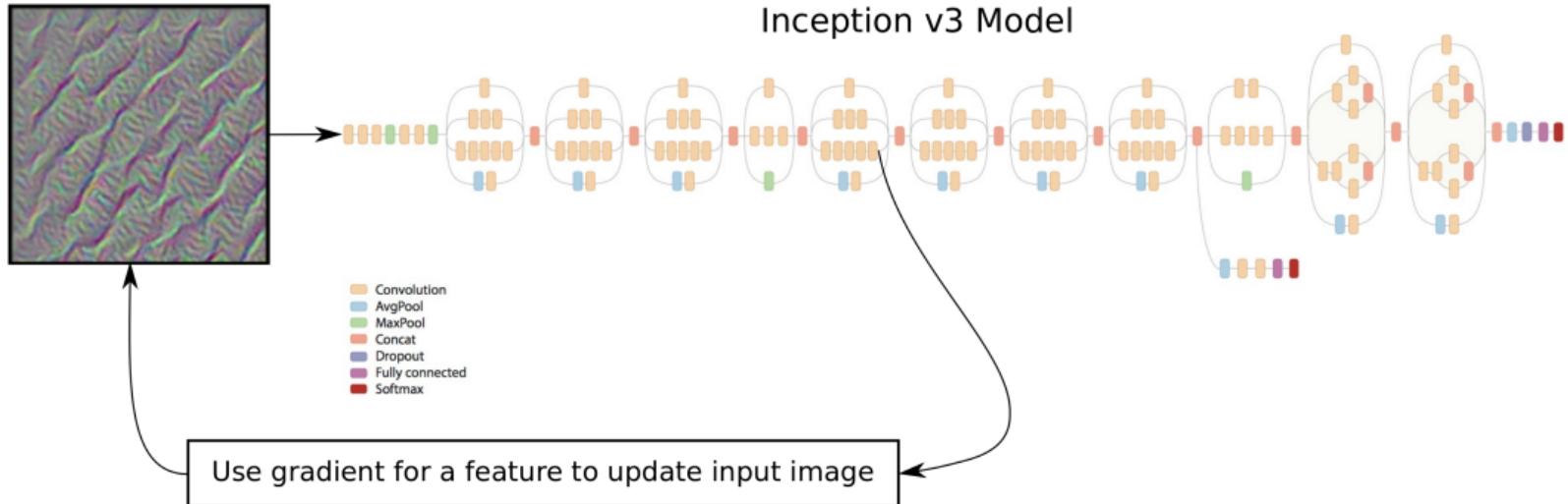


Layer 4

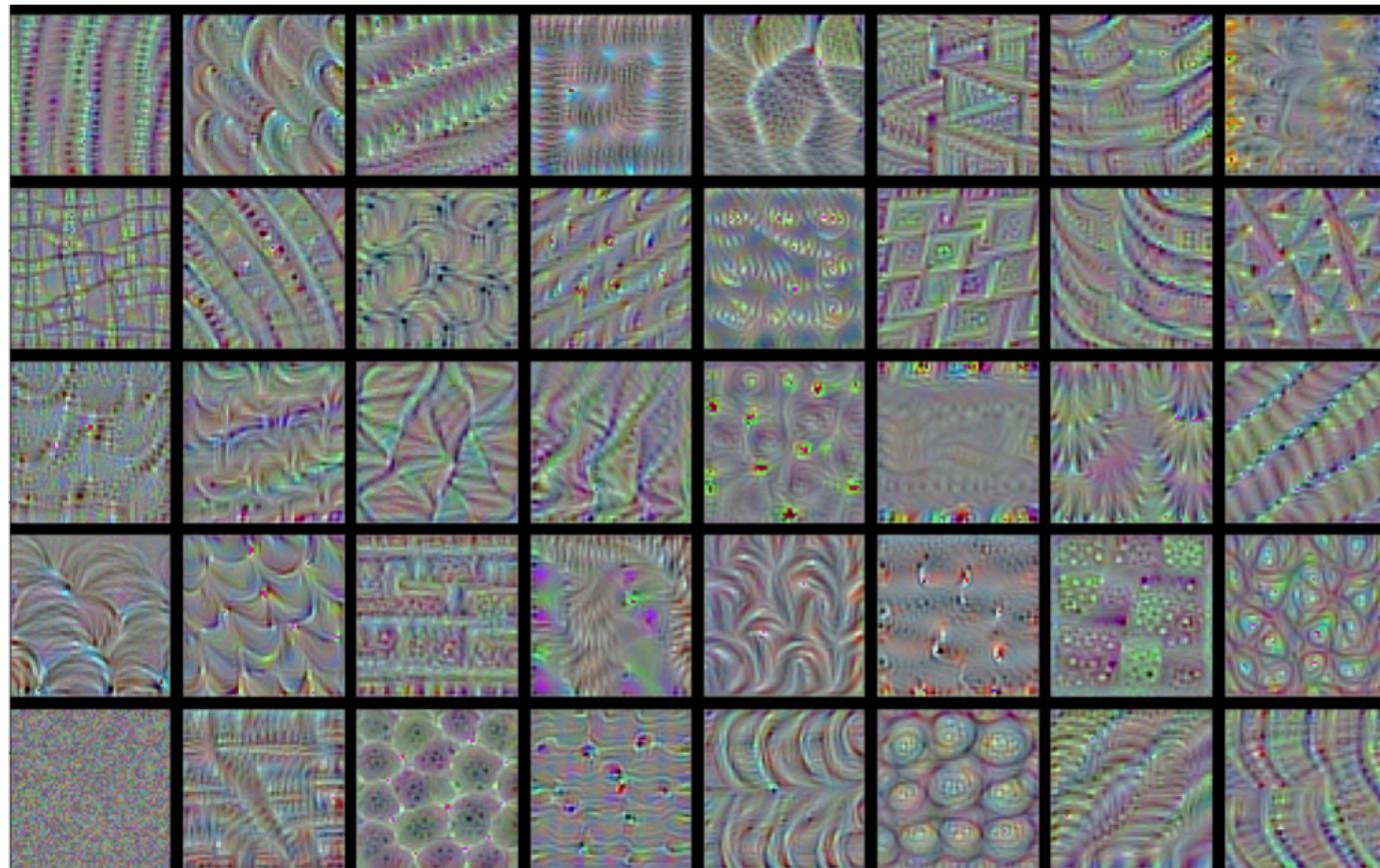


Layer 5

Есть несколько способов заглянуть внутрь свёрточной сетки, мы сейчас



Пример визуализации фильтров сверточных сетей



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>

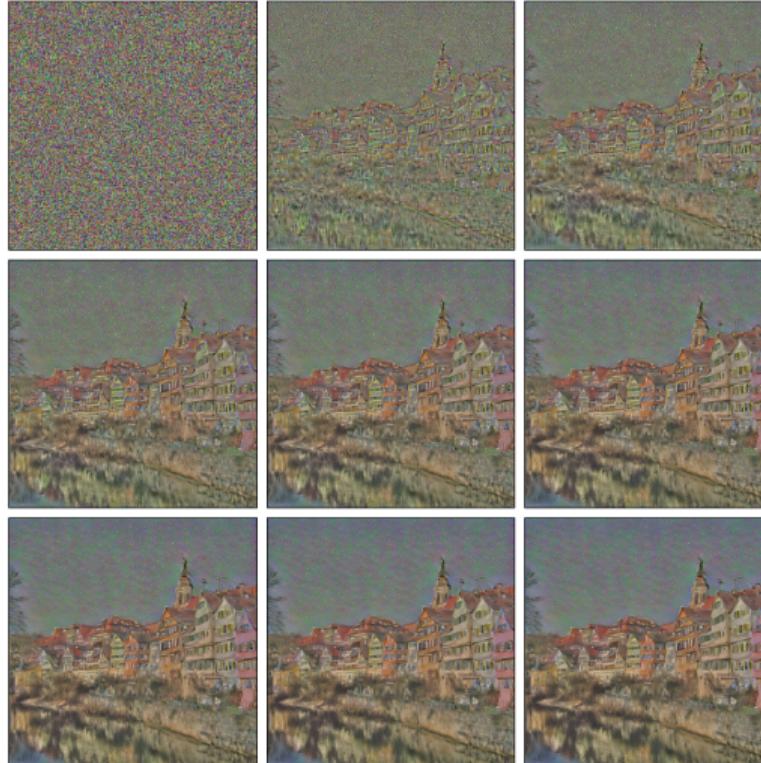
Переносим стиль!

Content loss



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

Content loss



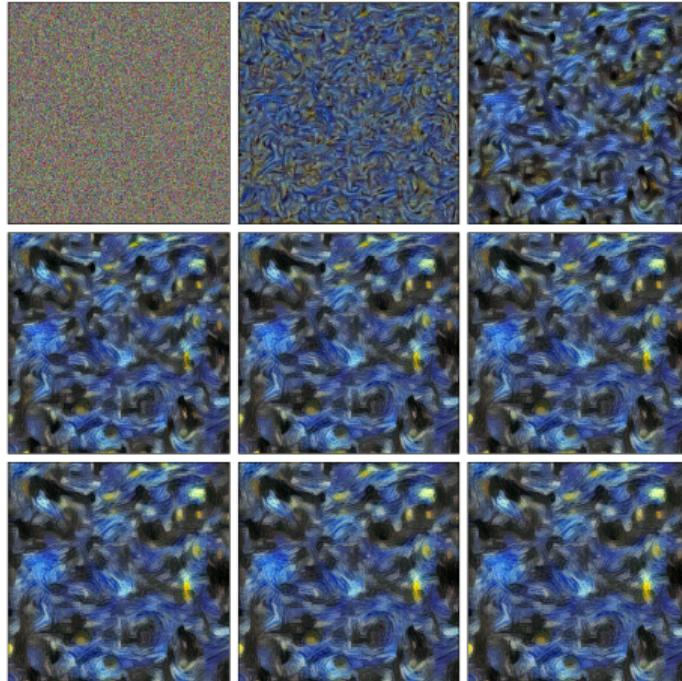
<https://habr.com/ru/company/mailru/blog/306916/>

Style loss



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

Style loss



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

Смесь нескольких стилей



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

Сегментация и локализация изображений

Сегментация и локализация

Semantic segmentation:



Object classification
+ localization:



water

duck

grass

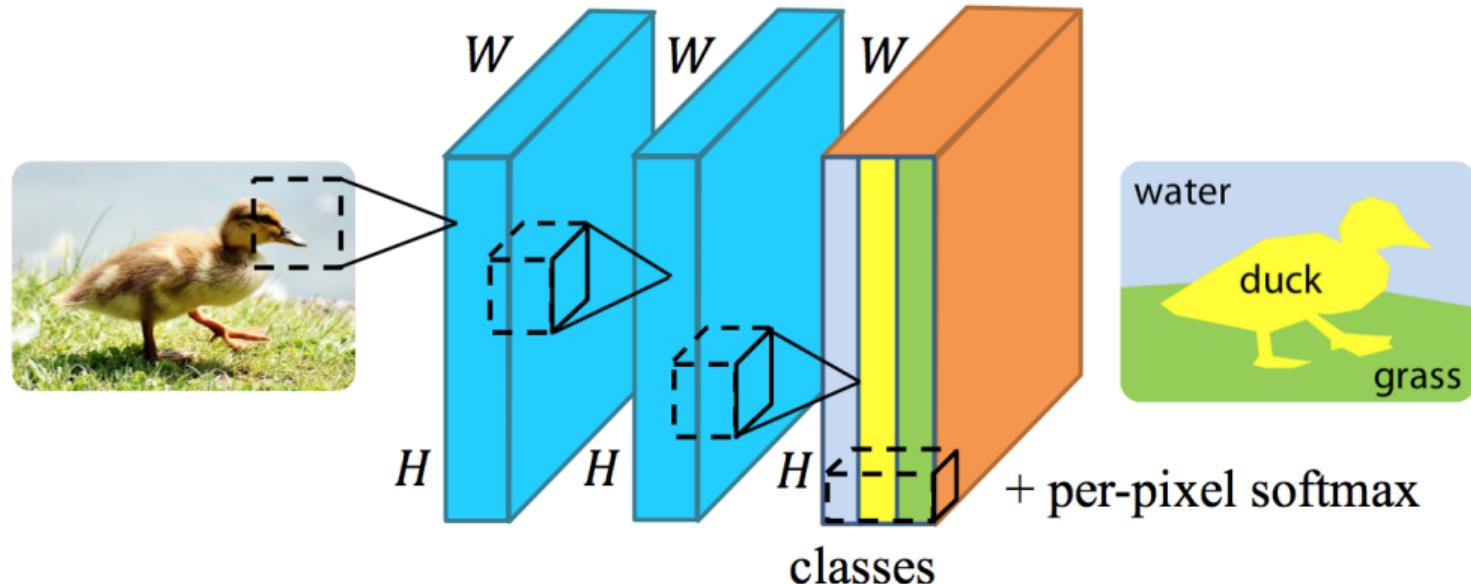
duck

Примеры



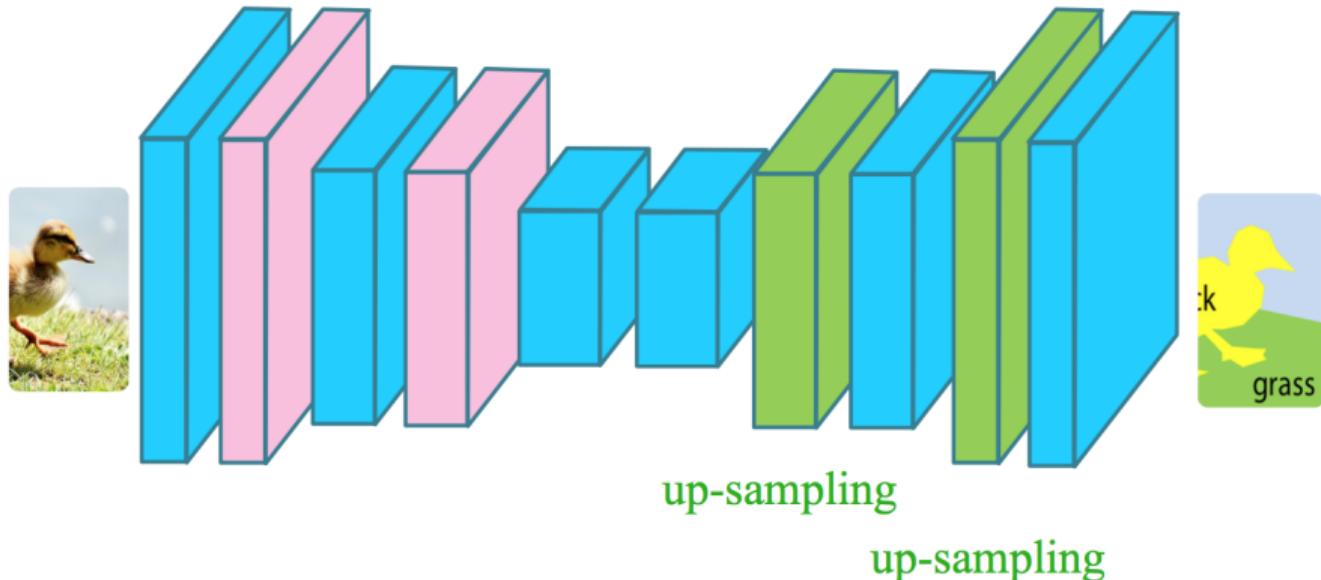
<https://www.youtube.com/watch?v=ZJMtDRbqH4o>

Сегментация



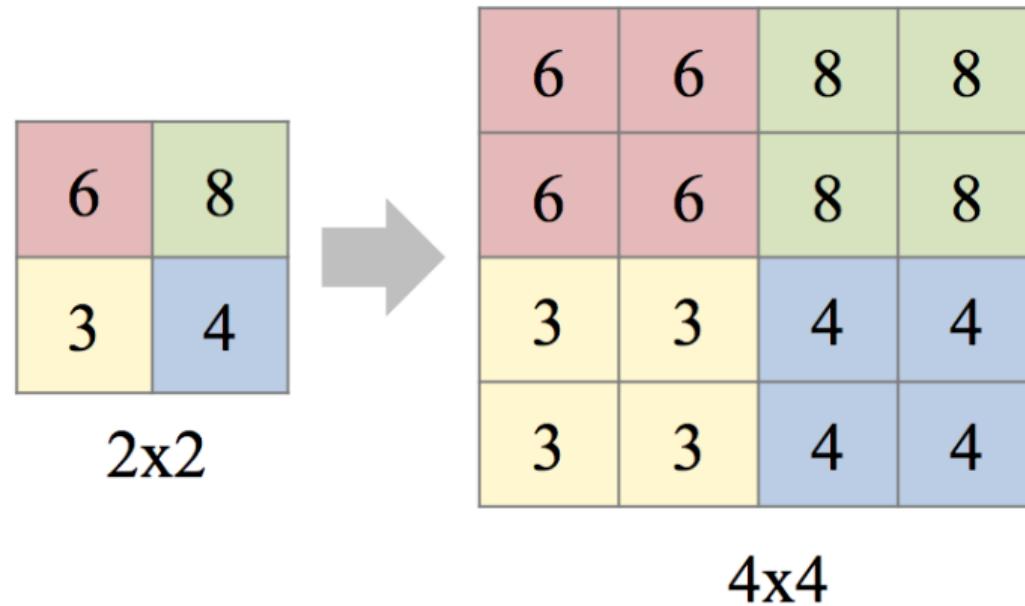
- Нам нужно научиться классифицировать каждый пиксель
- Куча свёрток и попиксельный softmax без пулинга (наивный подход)

Сегментация

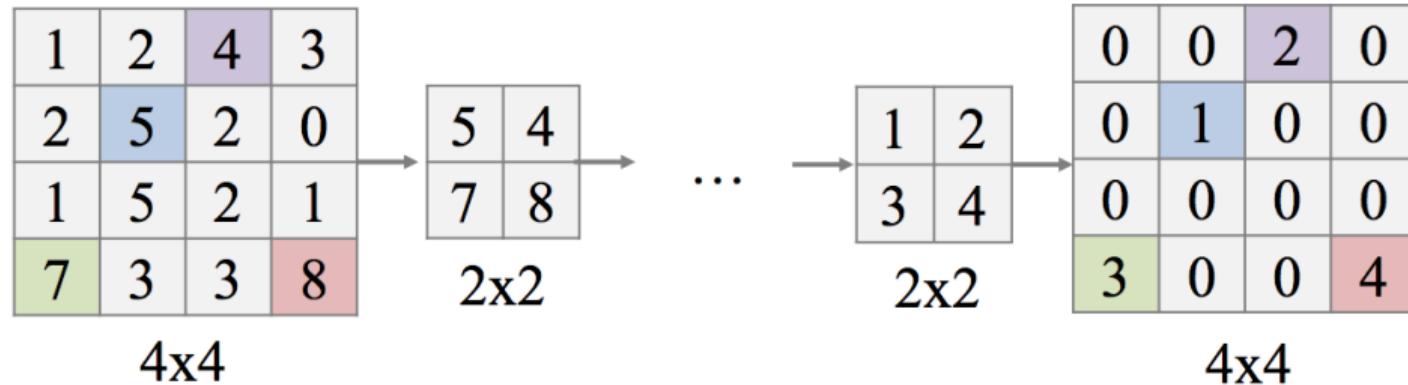


- Если захотим добавить пулинг, придётся делать анпулинг

Nearest neighbor unpooling

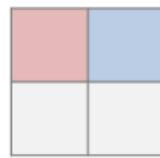


Max unpooling

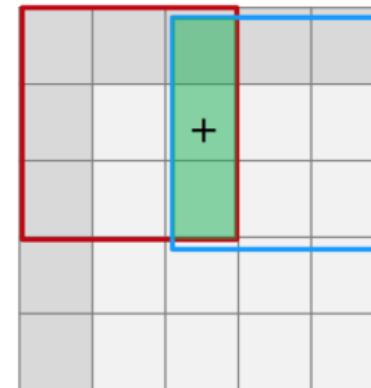


Learnable unpooling: Transpose convolution

Input: 2x2



Input gives
weight for
filter

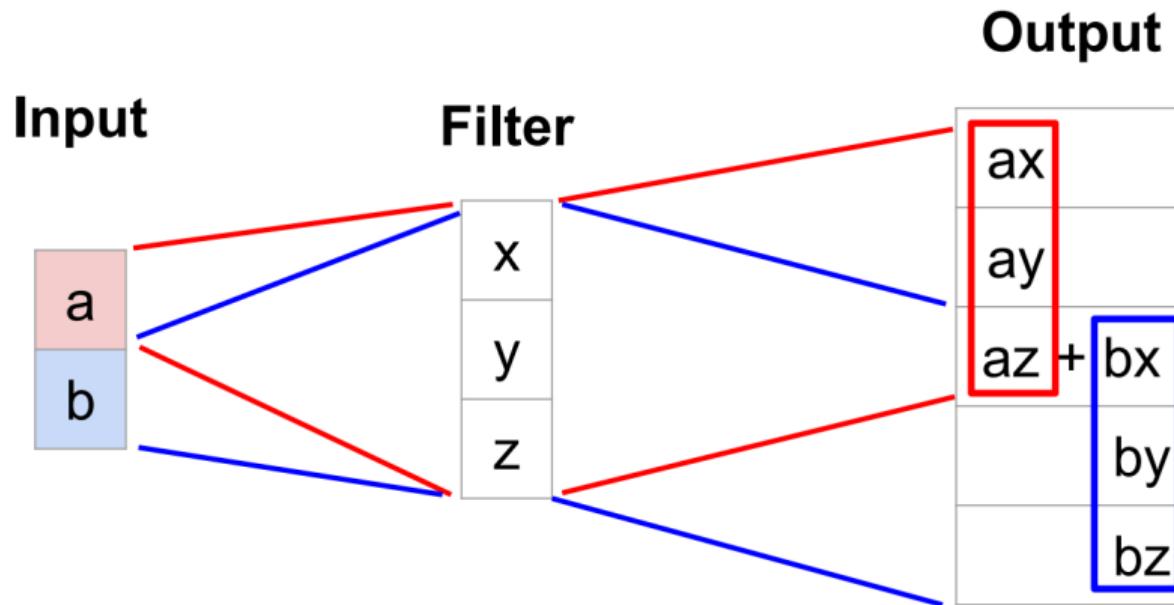


Stride: 2

Output: 4x4

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

Пример:



Почему называется Transpose convolution:

The diagram illustrates the computation of a transpose convolution result X from input vector \vec{a} and filter vector \vec{x} .

Input vector \vec{a} is represented as a column vector:

$$\vec{a} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

Filter vector \vec{x} is represented as a column vector:

$$\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The operation is labeled $*$ convolution.

The computation is shown as:

$$= \begin{bmatrix} x & y & z & 0 & 0 & 0 \\ 0 & x & y & z & 0 & 0 \\ 0 & 0 & x & y & z & 0 \\ 0 & 0 & 0 & x & y & z \end{bmatrix} \begin{bmatrix} 0 \\ a \\ b \\ c \\ d \\ 0 \end{bmatrix} = \begin{bmatrix} ay + bz \\ ax + by + cz \\ bx + cy + dz \\ cx + dy \end{bmatrix}$$

A large downward arrow points from the input vector \vec{a} to the equation $\vec{x} * \vec{a} = X\vec{a}$.

Source of a picture

Почему называется Transpose convolution:

$$\begin{bmatrix} x & 0 & 0 & 0 \\ y & x & 0 & 0 \\ z & y & x & 0 \\ 0 & z & y & x \\ 0 & 0 & z & y \\ 0 & 0 & 0 & z \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} ax \\ ay + bx \\ az + by + cx \\ bz + cy + dx \\ cz + dy \\ dz \end{bmatrix}$$

Source of a picture

Почему называется Transpose convolution:

$$\vec{x} * \vec{a} = X\vec{a}$$

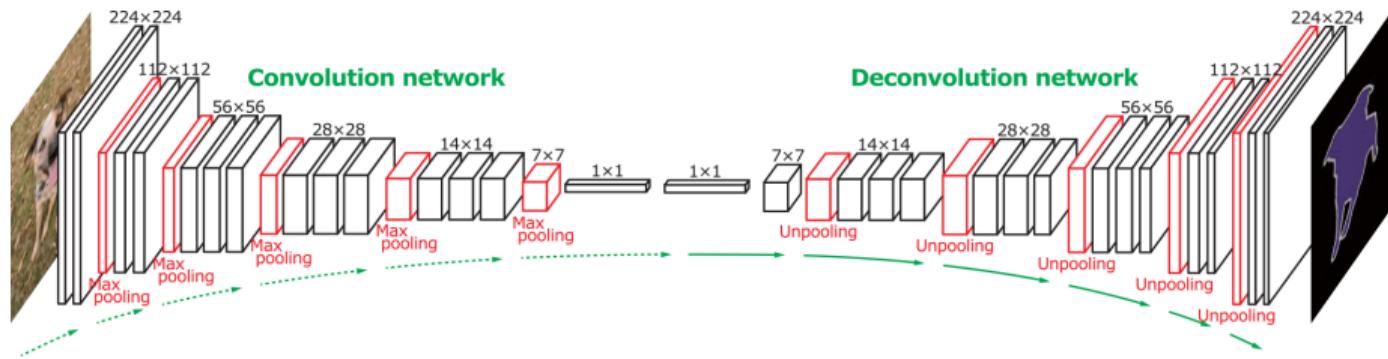
$$\begin{bmatrix} x & y & z & 0 & 0 & 0 \\ 0 & 0 & x & y & z & 0 \end{bmatrix} \begin{bmatrix} 0 \\ a \\ b \\ c \\ d \\ 0 \end{bmatrix} = \begin{bmatrix} ay + bz \\ bx + cy + dz \end{bmatrix}$$

$$\vec{x} *^T \vec{a} = X^T \vec{a}$$

$$\begin{bmatrix} x & 0 \\ y & 0 \\ z & x \\ 0 & y \\ 0 & z \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} ax \\ ay \\ az + bx \\ by \\ bz \\ 0 \end{bmatrix}$$

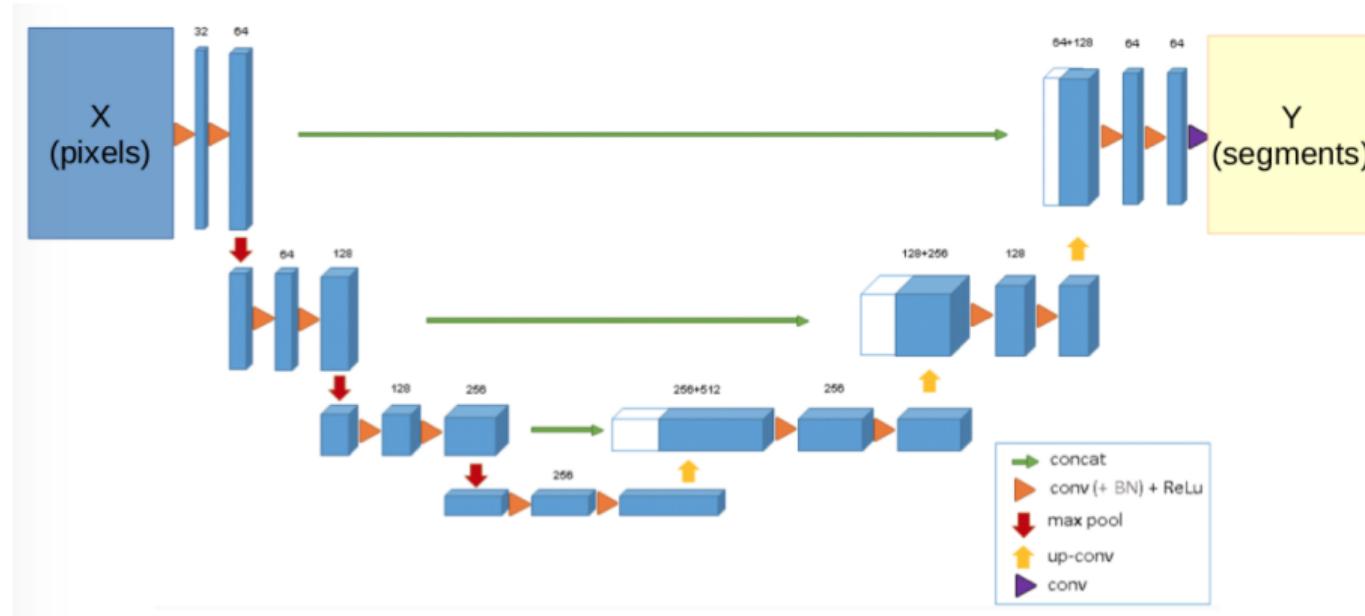
Source of a picture

Fully-convolution net



- Свернули в скрытое представление, развернули, спрогнозировали

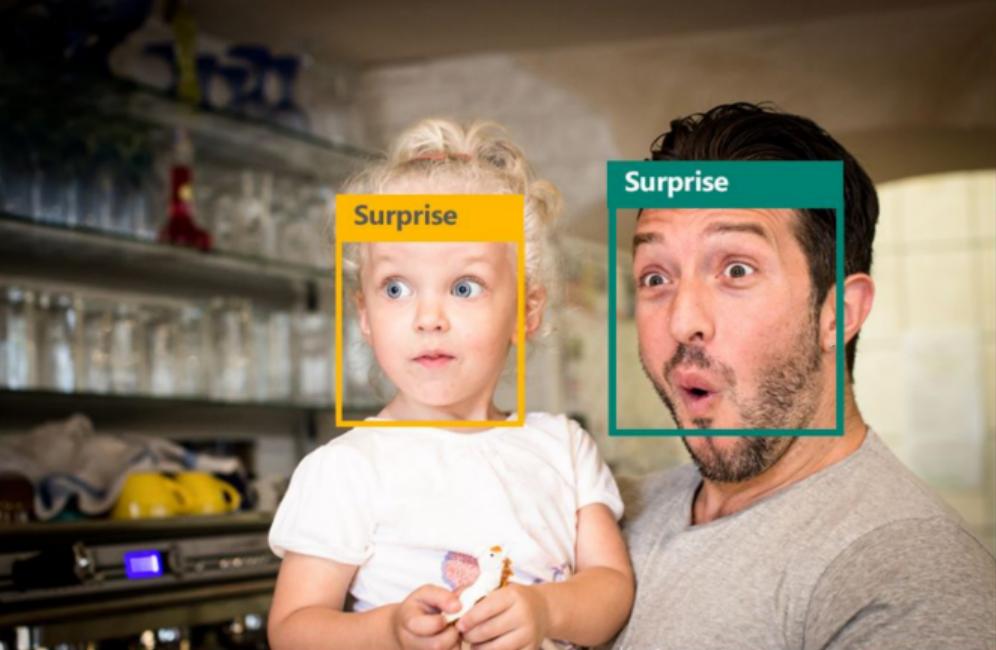
U-net



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

Локализация изображения

Примеры



The image shows a young child with blonde hair and a man with dark hair and a beard. Both individuals have wide-eyed expressions of surprise. The child's face is enclosed in a yellow bounding box, and the man's face is enclosed in a teal bounding box.

Surprise

Surprise

Neutral:  

Happiness:  

Surprise:  

Sadness:  

Anger:  

Disgust:  

Fear:  

Contempt:  



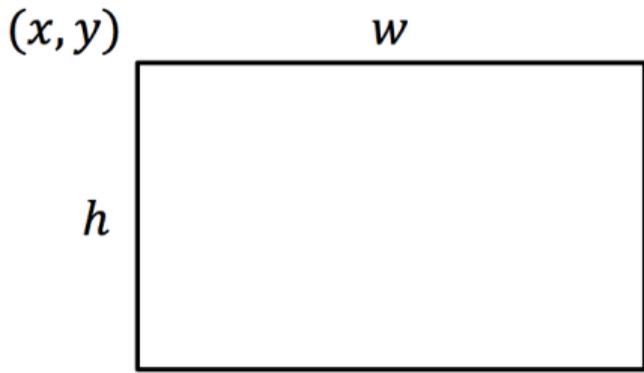
 Microsoft

Get started for free at projectoxford.ai

Примеры



Локализация



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

Локализация

