Сверточные нейронные сети для задачи сегментации

ЛИТПМ

МФТИ

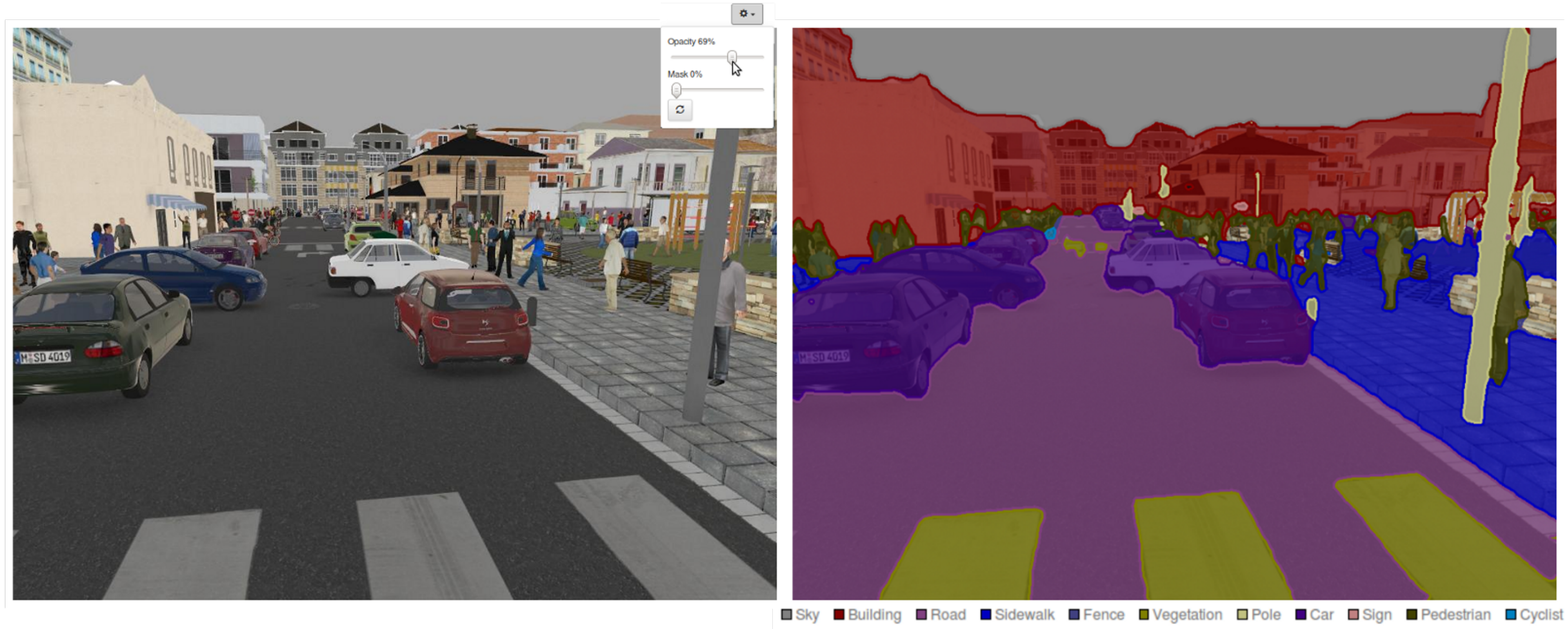
Кафедра технологий проектирования сложных технических систем

2022

Основные направления в CV

Сегментация

Решается задача классификации, но попиксельной

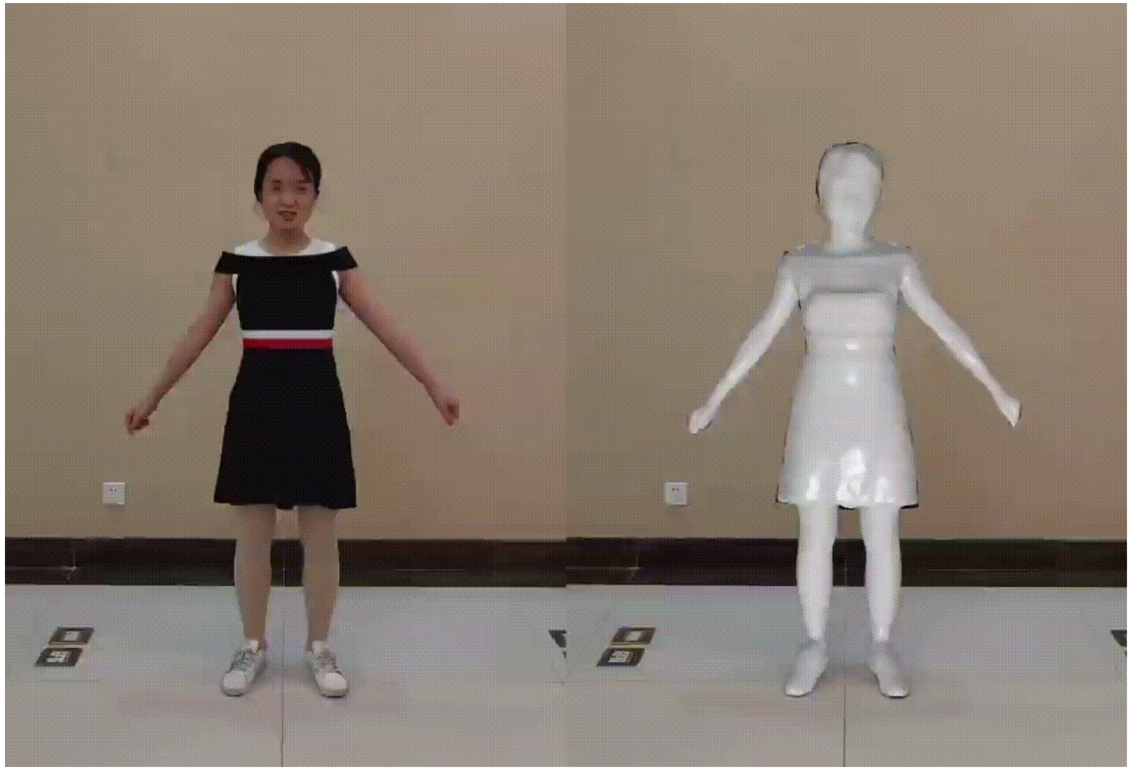


Изображение: ссылка на источник

Основные направления в CV

Построение 3D Объектов

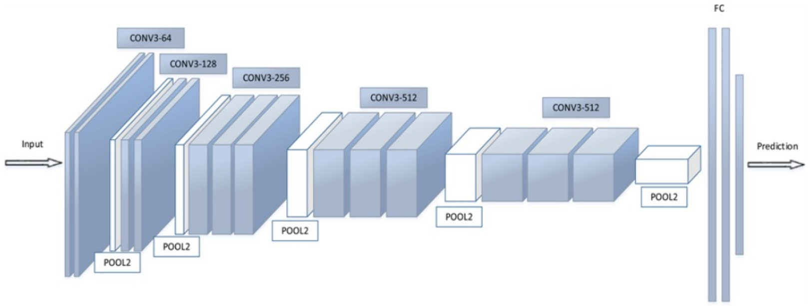
Решается задача построения объемного объекта (obj-файл) по двумерному изображению



Гифка: ссылка на источник, [Jiang et al. 2022]

Классификация

Классификация изображения

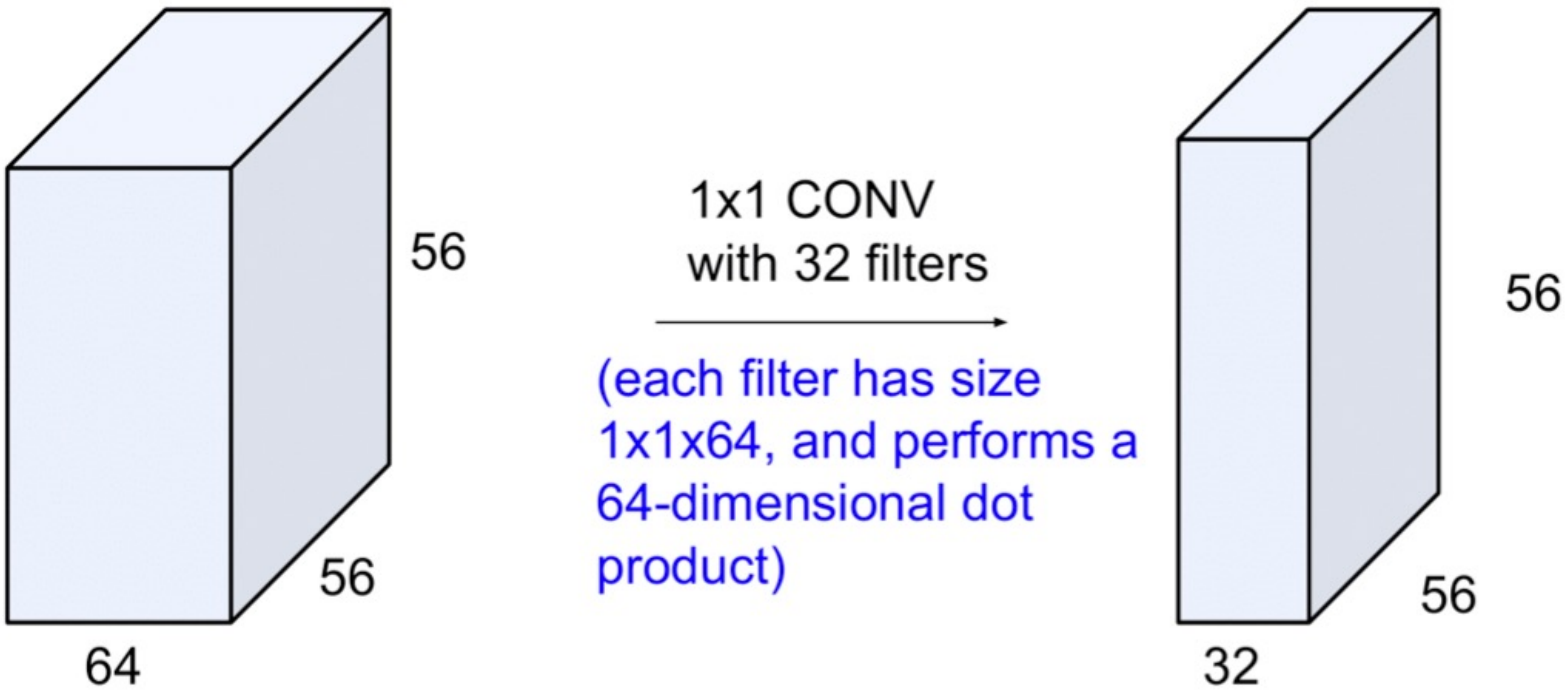


Cat (0.99)
Dog(0.01)

Классификация
пикселей изображения

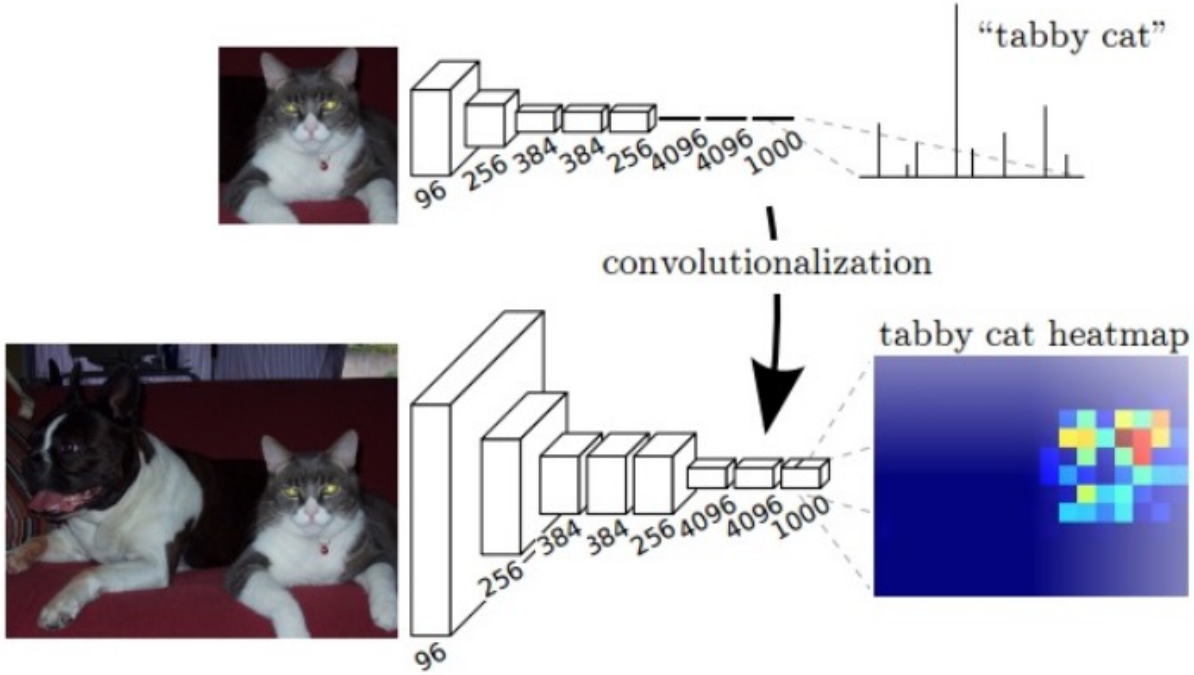
- Сегментация — попиксельная классификация
- Не требует большого количества данных
- Все сегментационные модели — это архитектуры вида FCN

Conv1x1 Recap



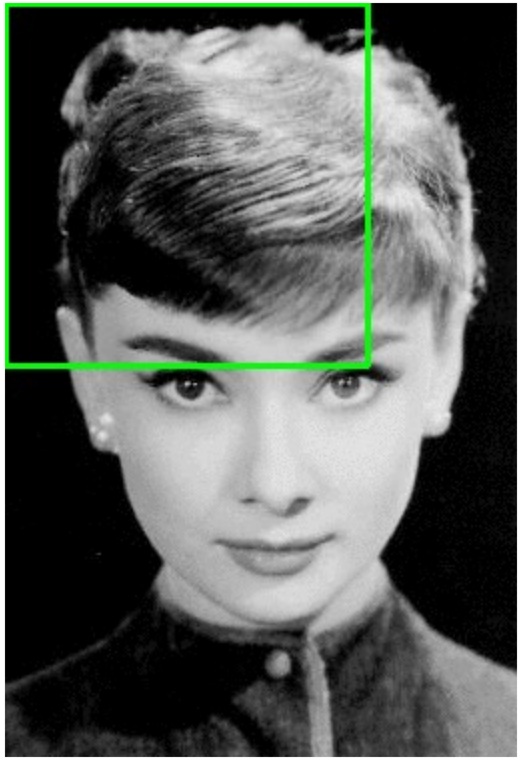
Source: Stanford CS231n Lecture 5 2017 by Fei-Fei Li & Justin Johnson & Serena Yeung

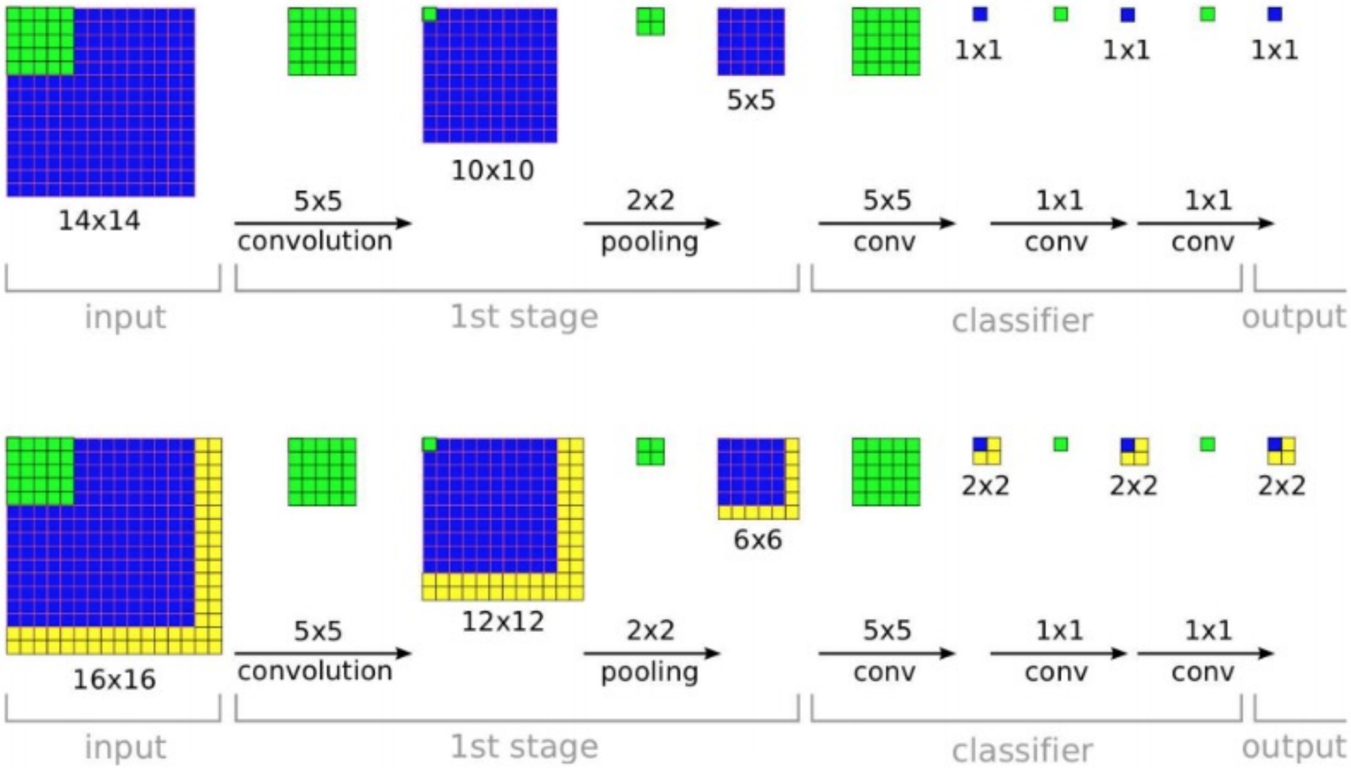
Fully Convolutional Network: FCN



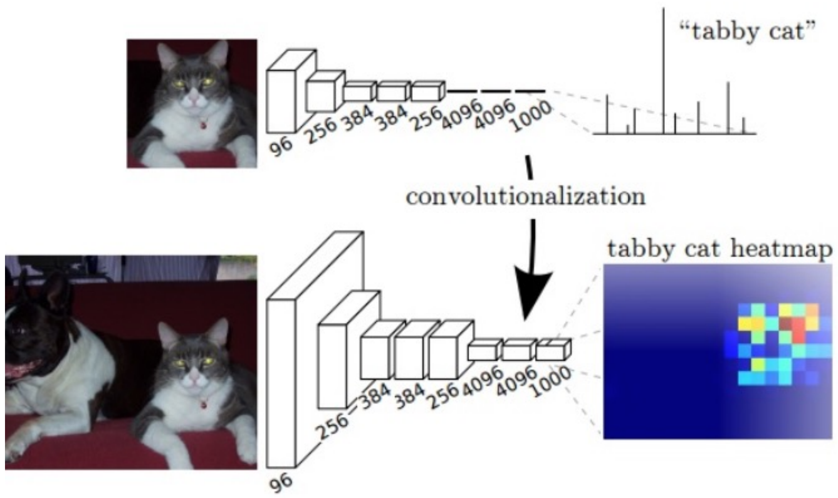
- Убираем FC и заменяем на Conv
- Получаем меньше обучаемых параметров
- На вход может получить изображение любого размера

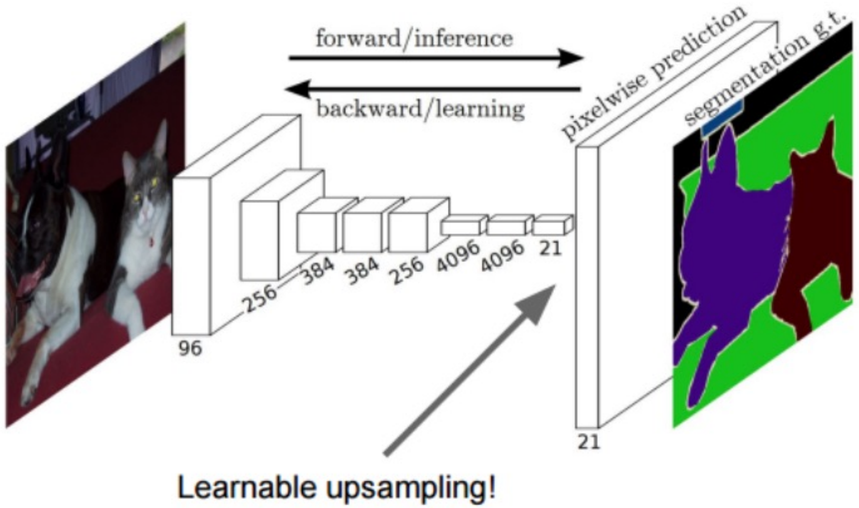
FCN = Efficient Sliding Window





Fully Convolutional Network: FCN

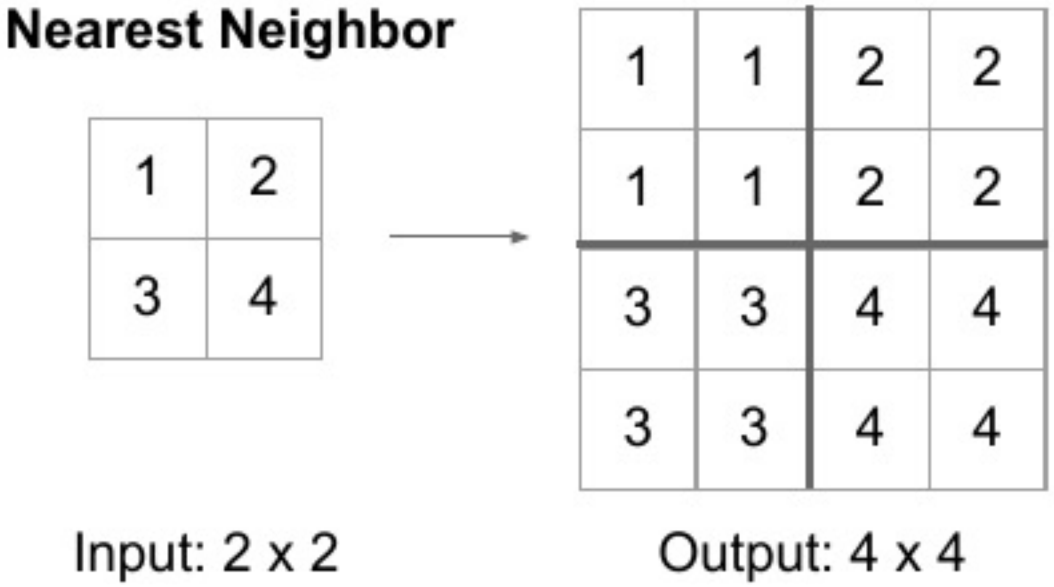




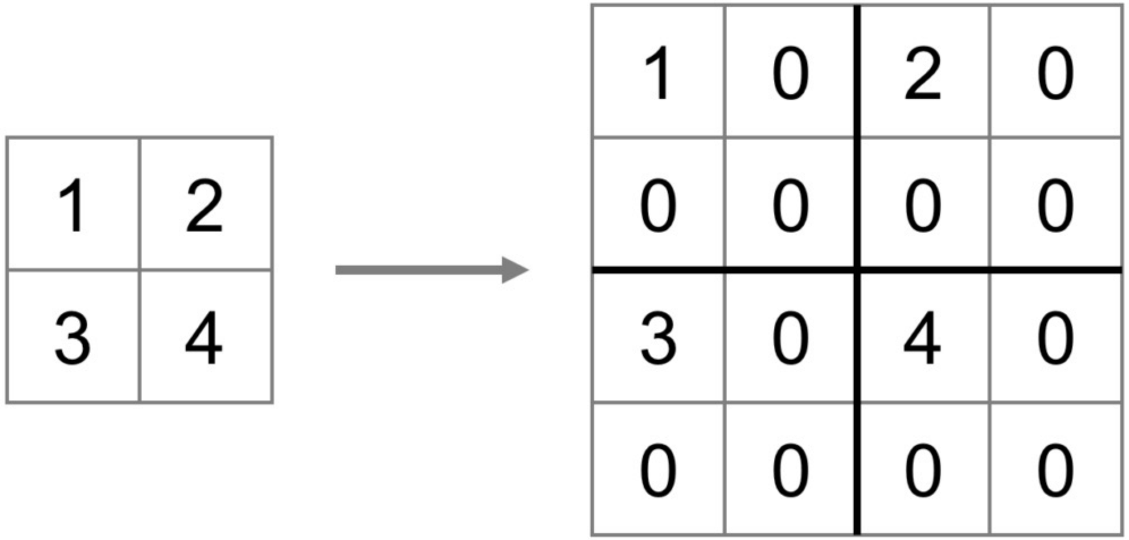
- Убираем FC и заменяем на Conv
- Добавляем декодер

Upsampling

Nearest neighbor

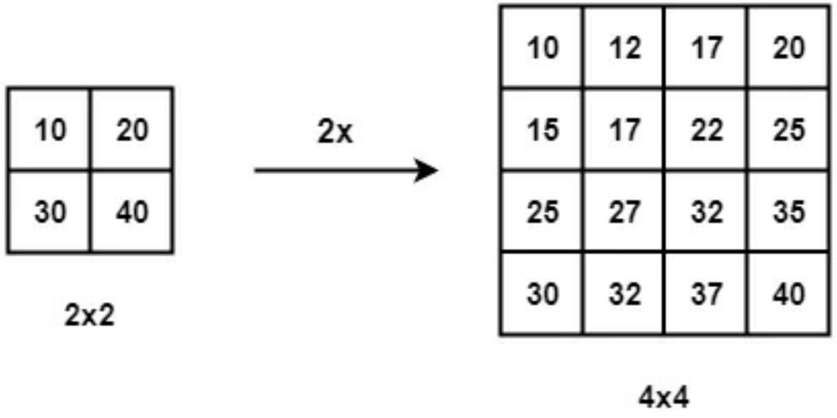


Bed of nails

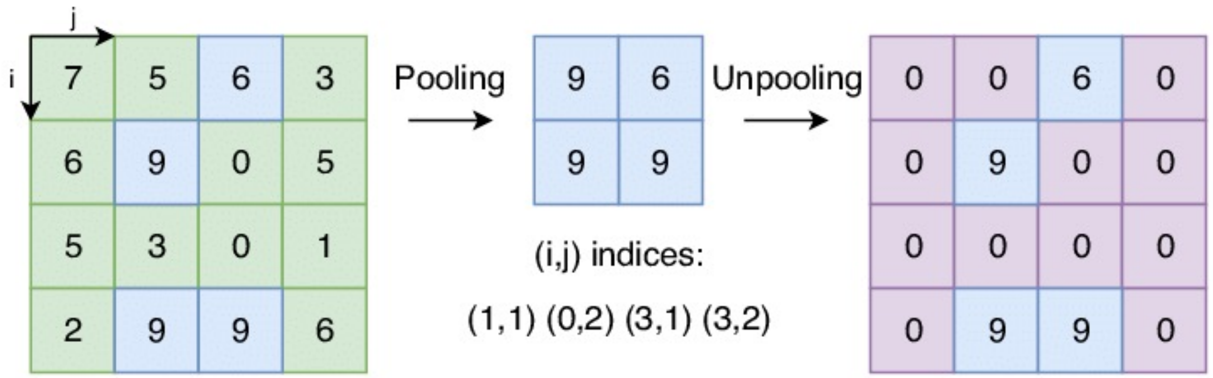


Upsampling

Bilinear interpolation

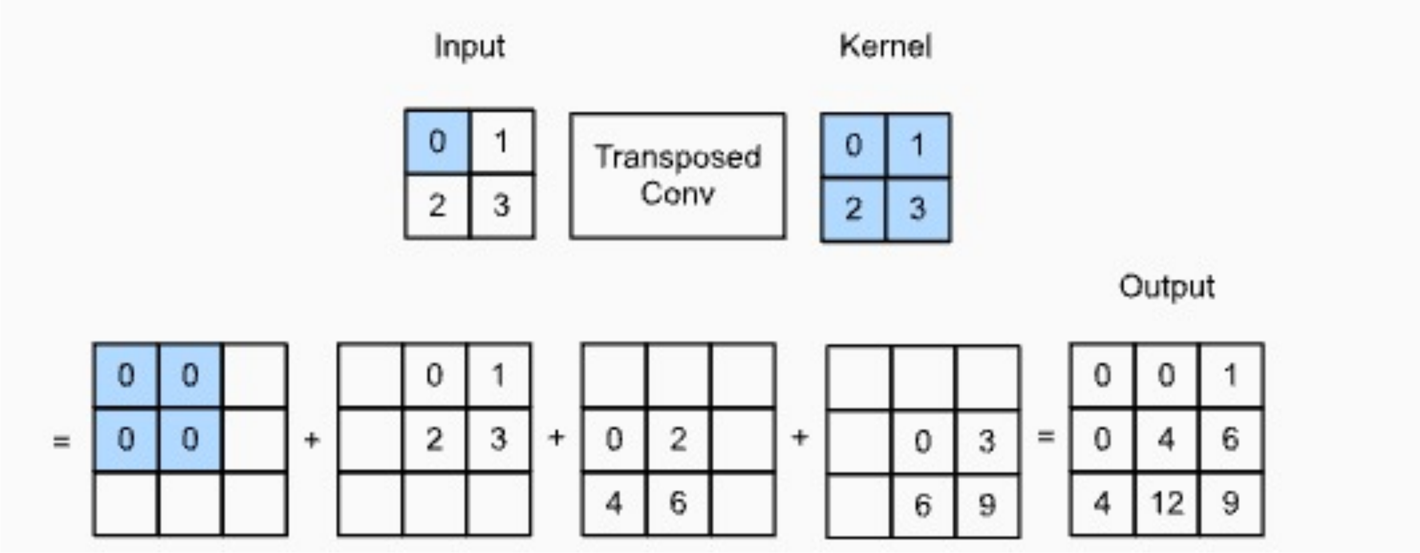


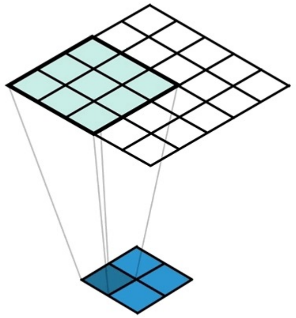
Max-Unpooling



Upsampling

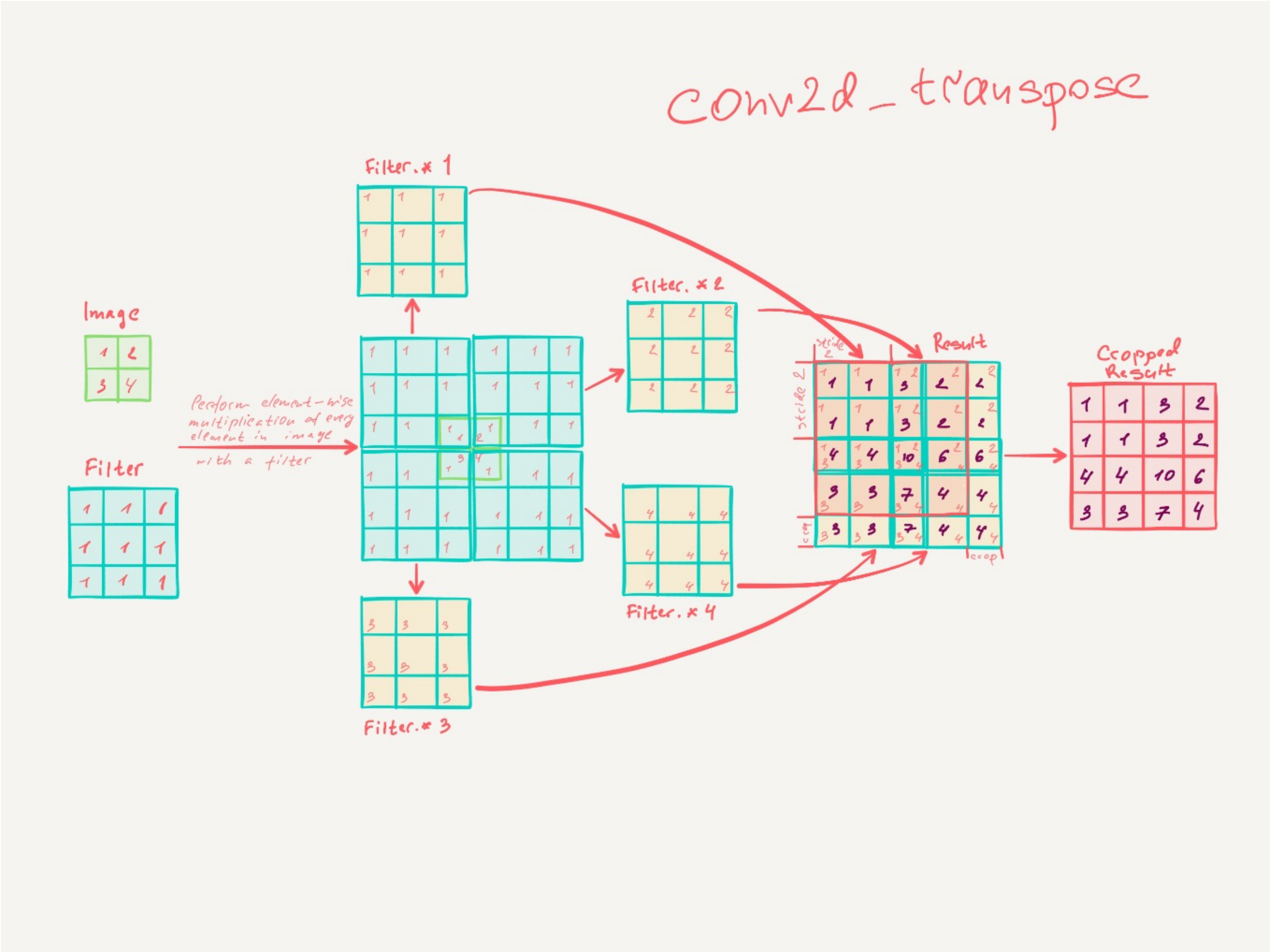
Transposed convolutions – also called fractionally strided convolutions – work by swapping the forward and backward passes of a convolution



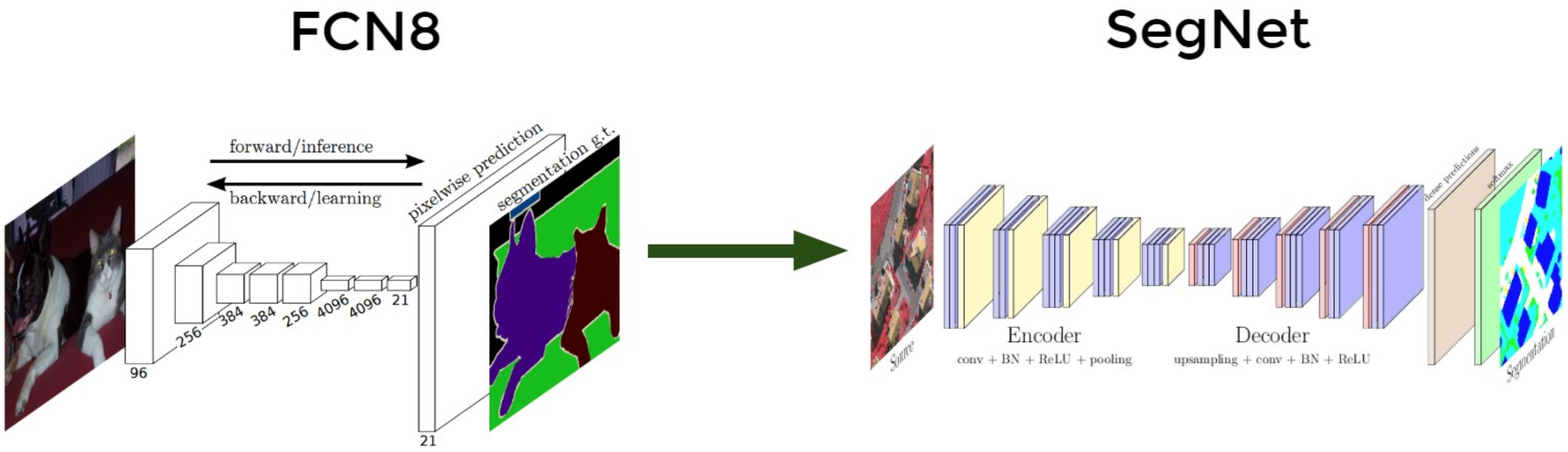


[Matthew D. Zeiler et al. CVPR 2010]

Upsampling


conv2d_transpose
Image
Filter
Perform element-wise multiplication of every element in image with a filter
Filter.* 1
Filter.* 2
Filter.* 3
Filter.* 4
Result
stride 2
crop
Cropped Result

FCN8 to SegNet

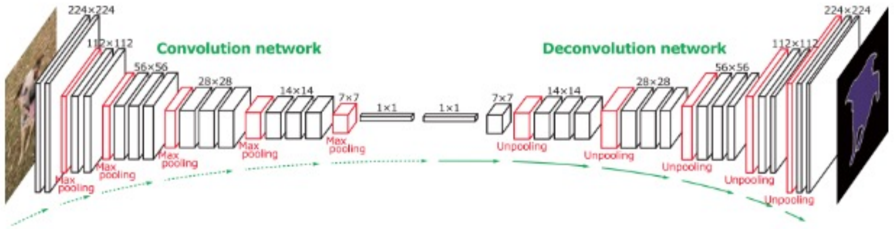


Заменить Upsampling на иерархический Upsampling

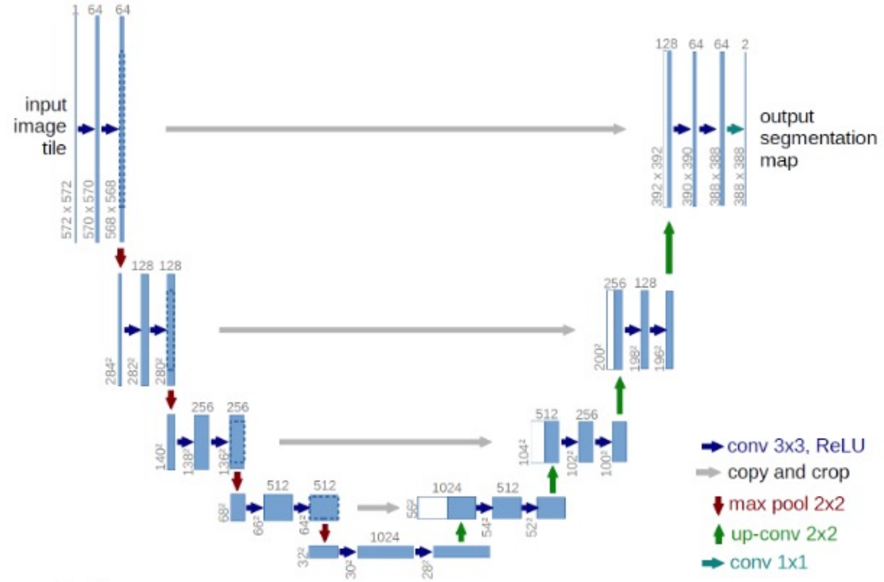
V. Badrinarayanan, A. Kendall, and R. Cipolla, “Segnet: A deep convolutional encoder-decoder architecture for image segmentation,” arXiv:1511.00561, 2015

Upsampling

SegNet



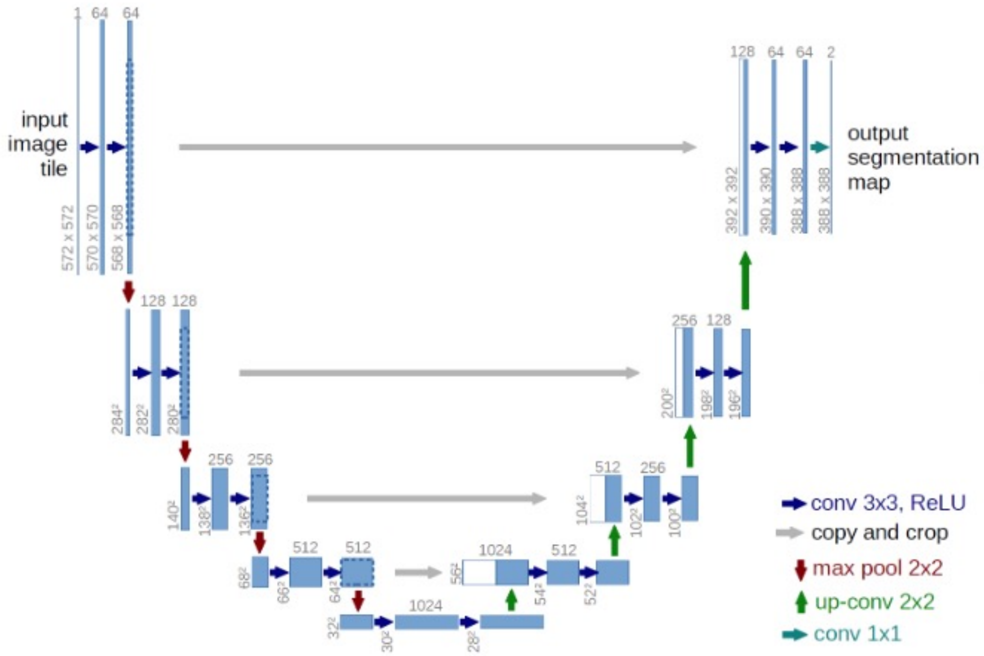
UNet



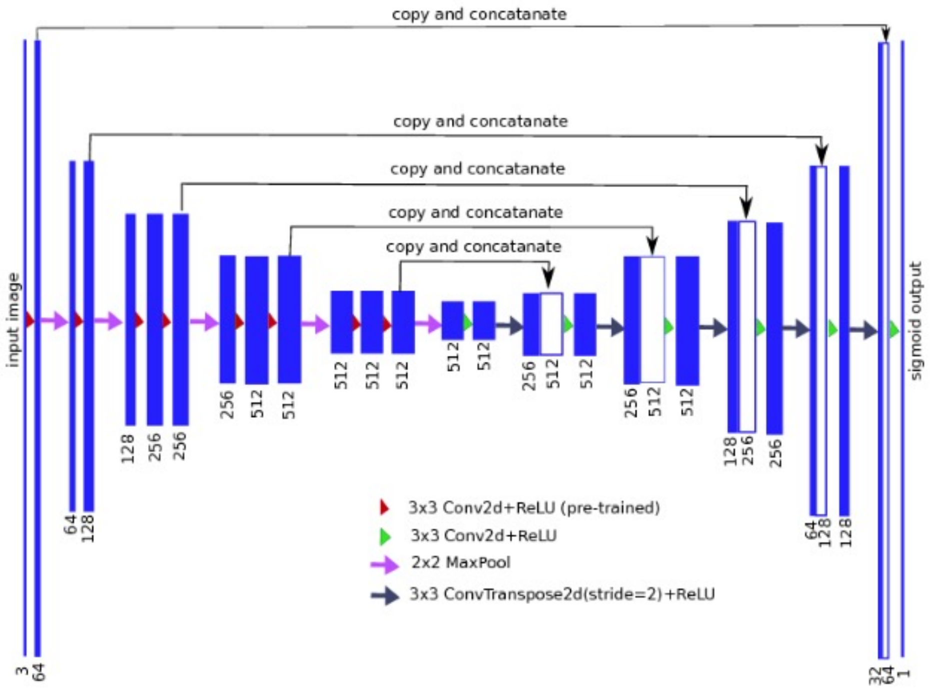
Added skip connections

O. Ronneberger P. Fischer T. Brox "U-net: Convolutional networks for biomedical image segmentation" Proc. Med. Image Comput. Comput.-Assisted Intervention pp. 234-241 2015.

Unet => TernausNet



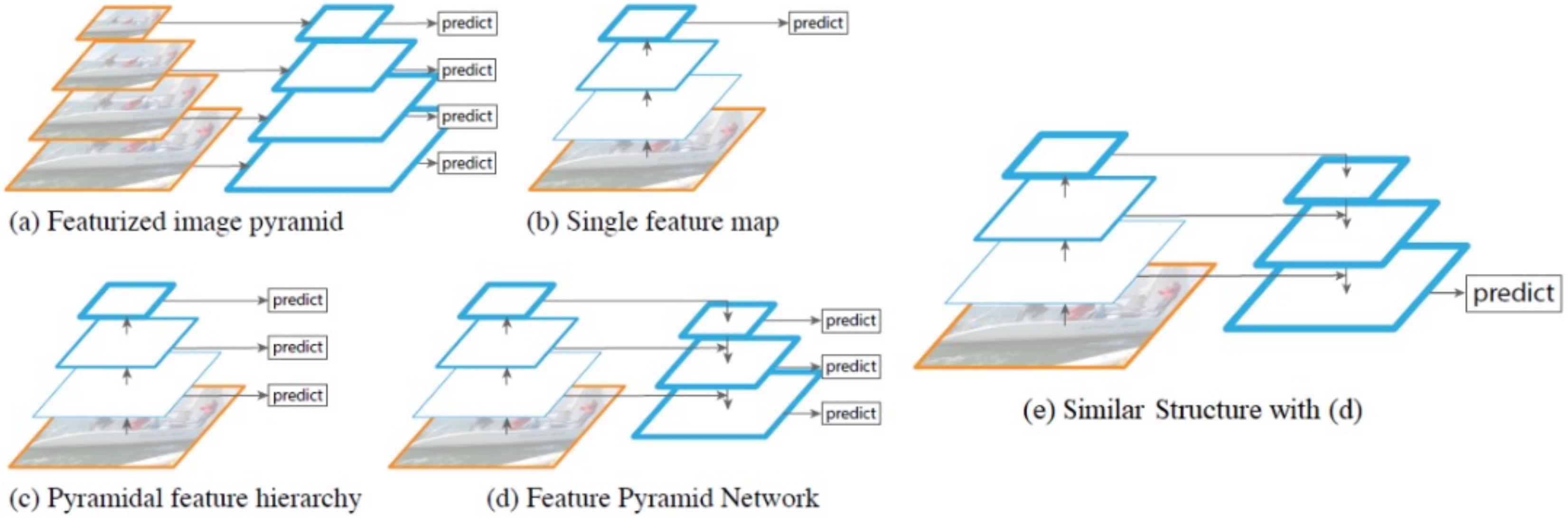
Text



Энкодер инициализируем весами с ImageNet

Iglovikov, V., Shvets, A.: Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. arXiv preprint arXiv:1801.05746 (2018)

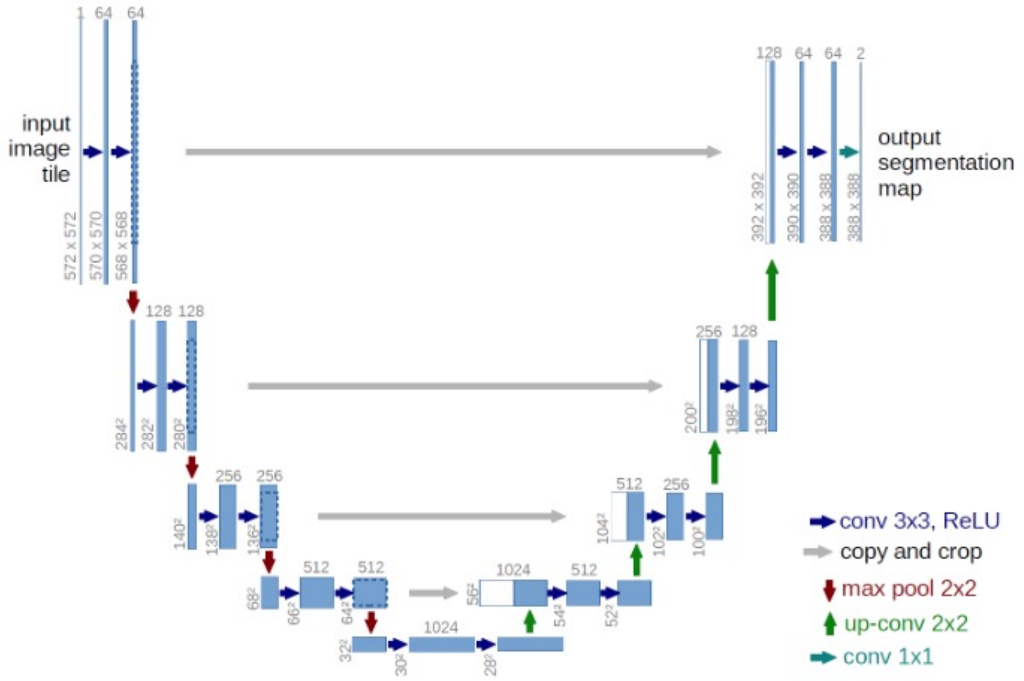
Feature Pyramid Networks (FPN)

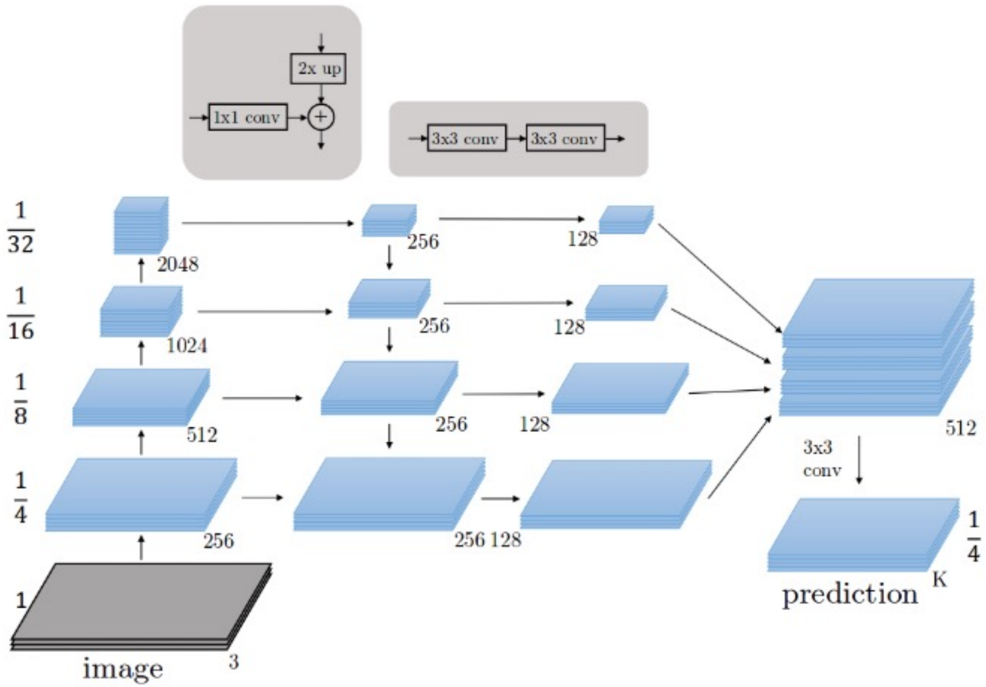


1. Легко добавить во многие архитектуры.
2. Помогает с multiscale

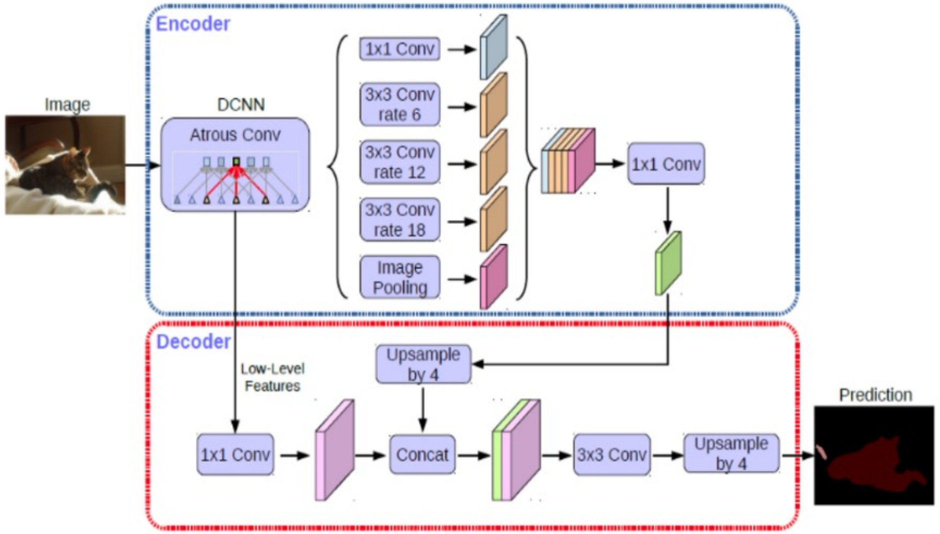
Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie; The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2117-2125

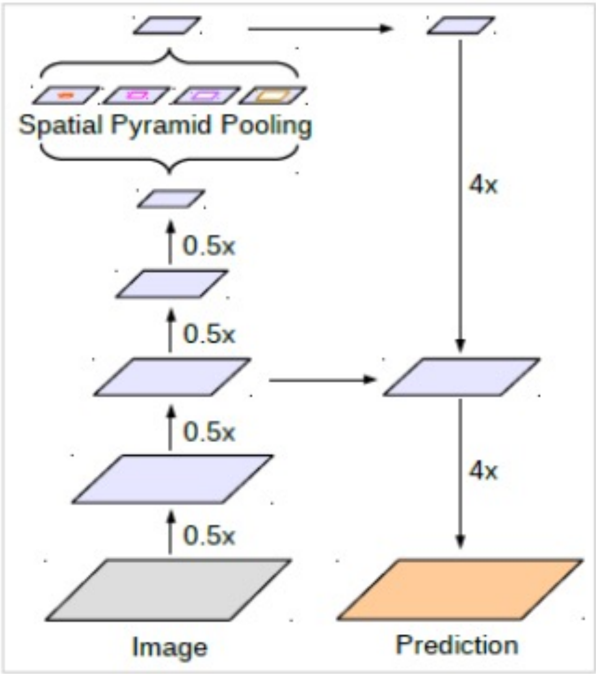
Unet + FPN

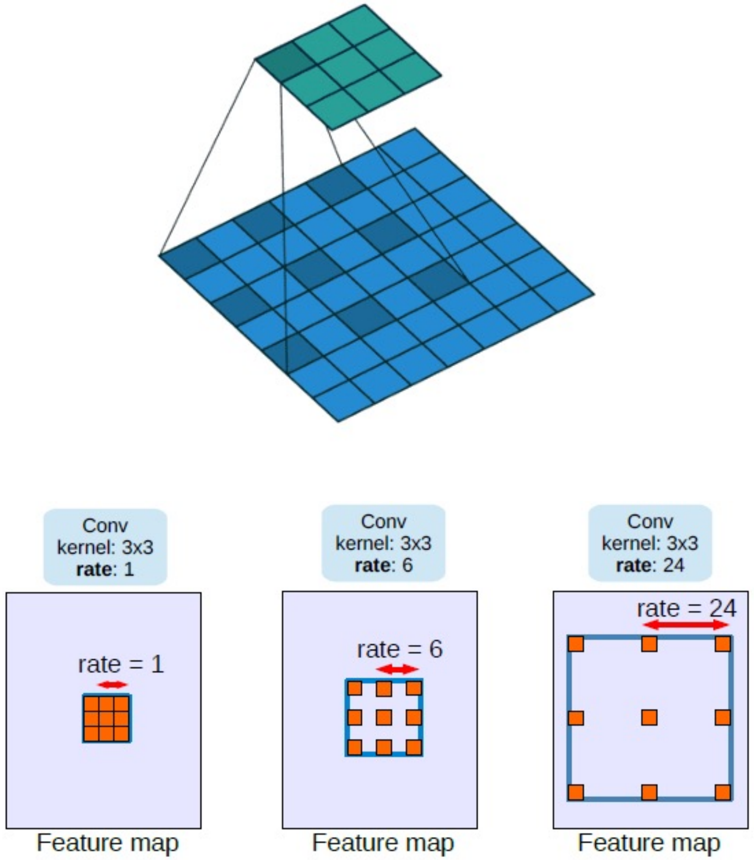

1 64 64
input image tile
572 x 572
570 x 570
568 x 568
128 128
284²
282²
280²
256 256
140²
138²
136²
512 512
68²
66²
64²
1024
32²
30²
28²
1024 512
56²
54²
52²
512 256
104²
102²
100²
256 128
200²
198²
196²
128 64 64 2
392 x 392
390 x 390
388 x 388
388 x 388
output segmentation map
conv 3x3, ReLU
copy and crop
max pool 2x2
up-conv 2x2
conv 1x1


2x up
1x1 conv
3x3 conv
3x3 conv
1/32
1/16
1/8
1/4
1
2048
1024
512
256
256
256
256
256 128
128
128
128
512
3x3 conv
1/4
K
prediction
image
3

DeepLabV3







Atrous Separable Convolution

Segmentation Loss Function

Каждый пиксель классификатор => Categorical / Binary Cross Entropy(CCE, BCE)

Но! Метрика Dice / Jaccard

Dice / Jaccard недифференцируемы =>

Soft Dice / Soft Jaccard

и добавляем в loss

Lovasz-Softmax loss

Использовать для FineTune

$$CCE = \sum_c p(x) \log q(x)$$

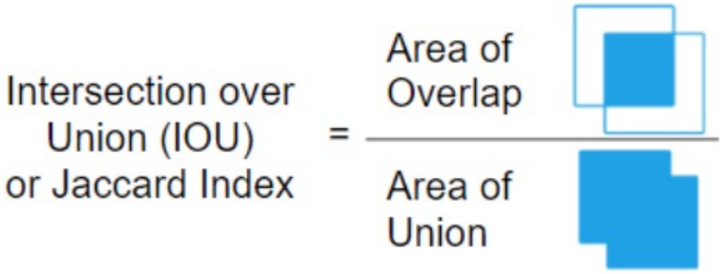
$$LOSS = BCE - \ln(DICE)$$

$$BCE = -\sum_i \left(y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)\right)$$

$$DICE = 2\frac{\sum_i y_i p_i}{\sum_u y_i + \sum p_i}$$

Berman, M., Rannen Triki, A., Blaschko, M.B.: The lovász-softmax loss: a tractable surrogate for the optimization of the intersection-over-union measure in neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4413–4421 (2018)

Segmentation: Metric



Чаще всего используют Dice - особенно в медицинских снимках и Jaccard (IoU)

Table 1. The three similarity coefficients

Similarity Coefficient (X,Y)	Actual Formula
Dice Coefficient	$2\frac{\lvert X \cap Y \rvert}{\lvert X \rvert + \lvert Y \rvert}$
Cosine Coefficient	$\frac{\lvert X \cap Y \rvert}{\lvert X \rvert^{1/2} . \lvert Y \rvert^{1/2}}$
Jaccard Coefficient	$\frac{\lvert X \cap Y \rvert}{\lvert X \rvert + \lvert Y \rvert - \lvert X \cap Y \rvert}$