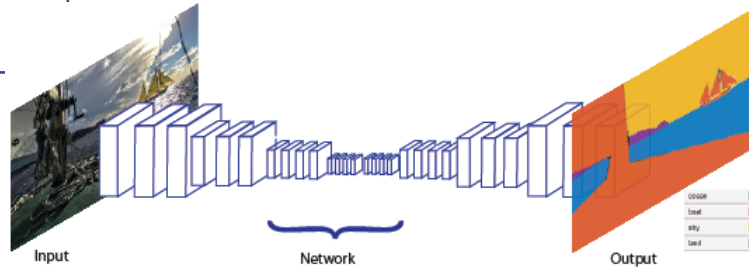
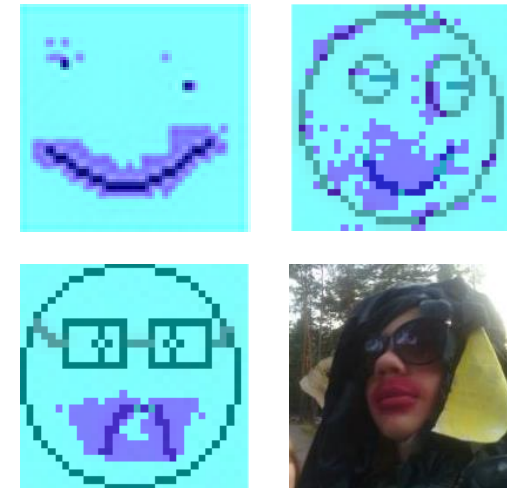
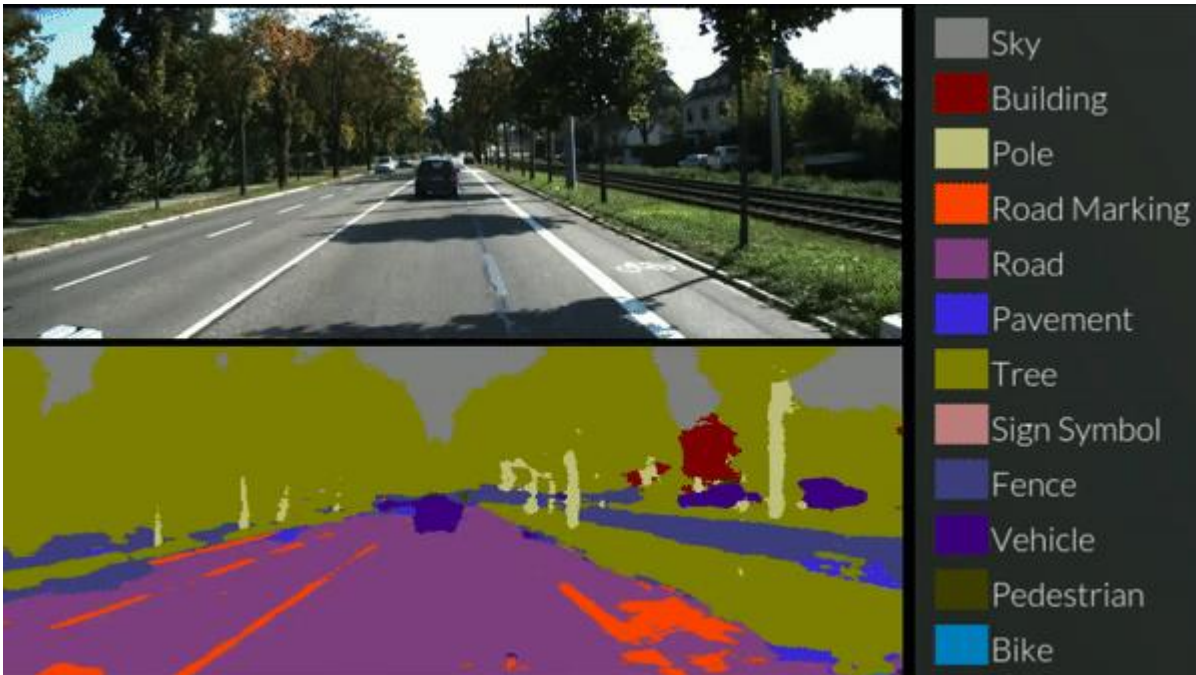
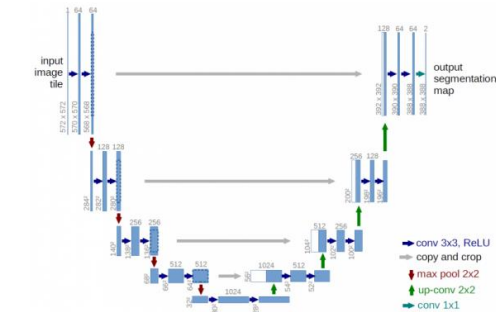


Сегментация как случай классификации / Semantic segmentation is a particular case of a classification problem

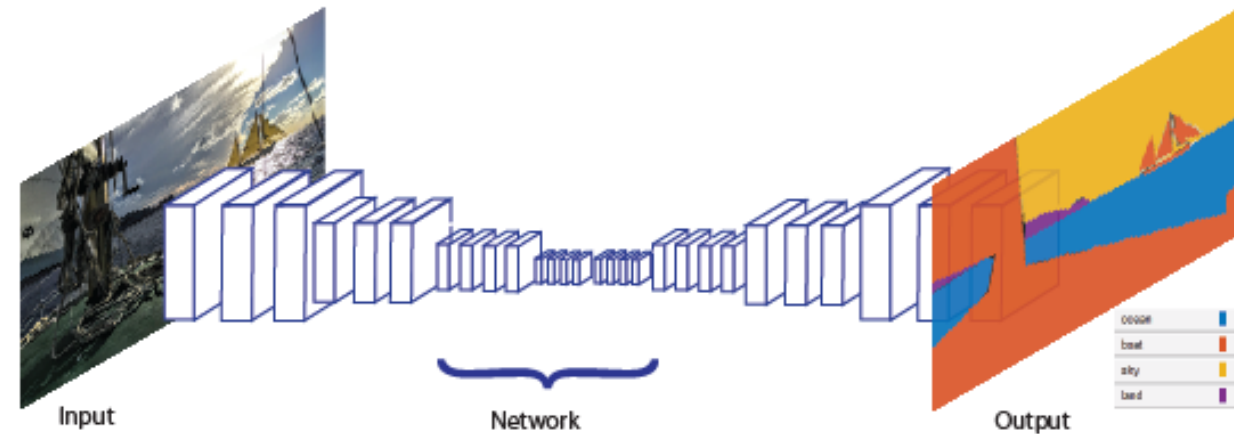
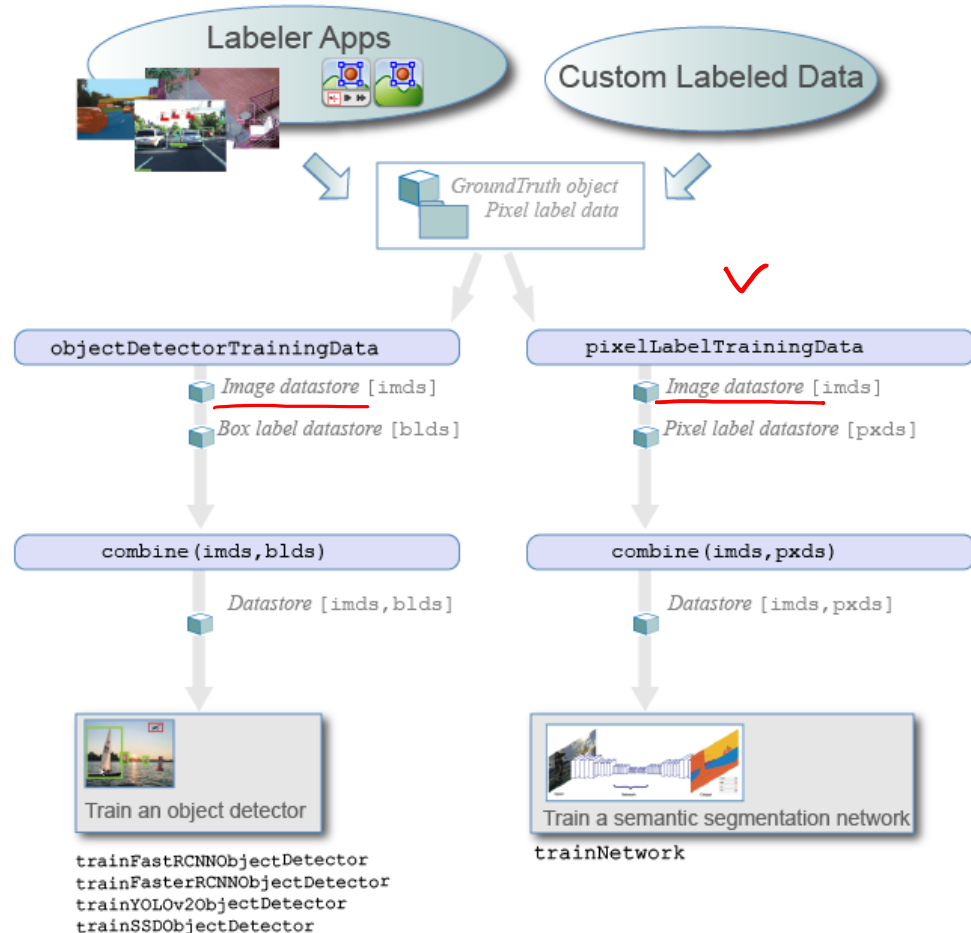


Архитектура U-Net



Суть / The main idea

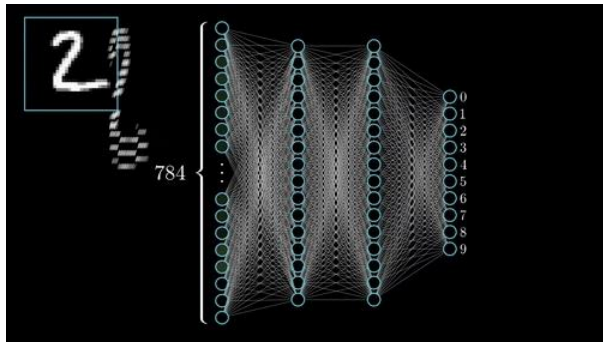
Семантическая сегментация — разбиение изображения на объекты с определением типов этих объектов [habr.com]. При этом каждому пикселю на изображении ставится в соответствие класс. То есть семантическая сегментация в машинном обучении – это частный случай классификации.



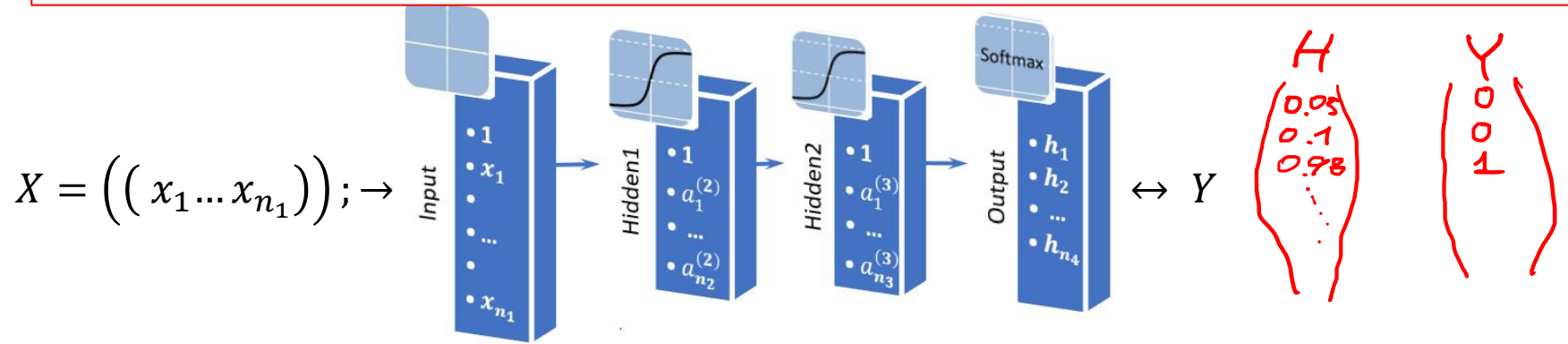
Классификация с помощью полносвязной сети / Solving classification problem using fully connected neural network

Характерная архитектура сети. Количество слоев - l ; количество нейронов в k -ом слое - n_k (n_l - количество нейронов в выходном слое); логистическая функция активации в скрытых слоях и функция активации «софтмакс» в выходном слое.

$$J(\Theta^{(k)}) = -\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^{n_l} (y_j^{(i)} \ln(h_j^{(i)}) + (1 - y_j^{(i)}) \ln(1 - h_j^{(i)})) + \frac{\lambda}{2m} \sum_{k=1}^{l-1} \sum_{i=1}^{n_k} \sum_{j=1}^{n_{k+1}} (\theta_{ij}^{(k)})^2 \Rightarrow \min.$$



Animation from "3Blue1Brown"



$$X = \begin{pmatrix} x_1 & \dots & x_{n_1} \end{pmatrix}; \rightarrow$$

$$\leftrightarrow Y$$

$$l = 4$$

$$n_l = 10$$

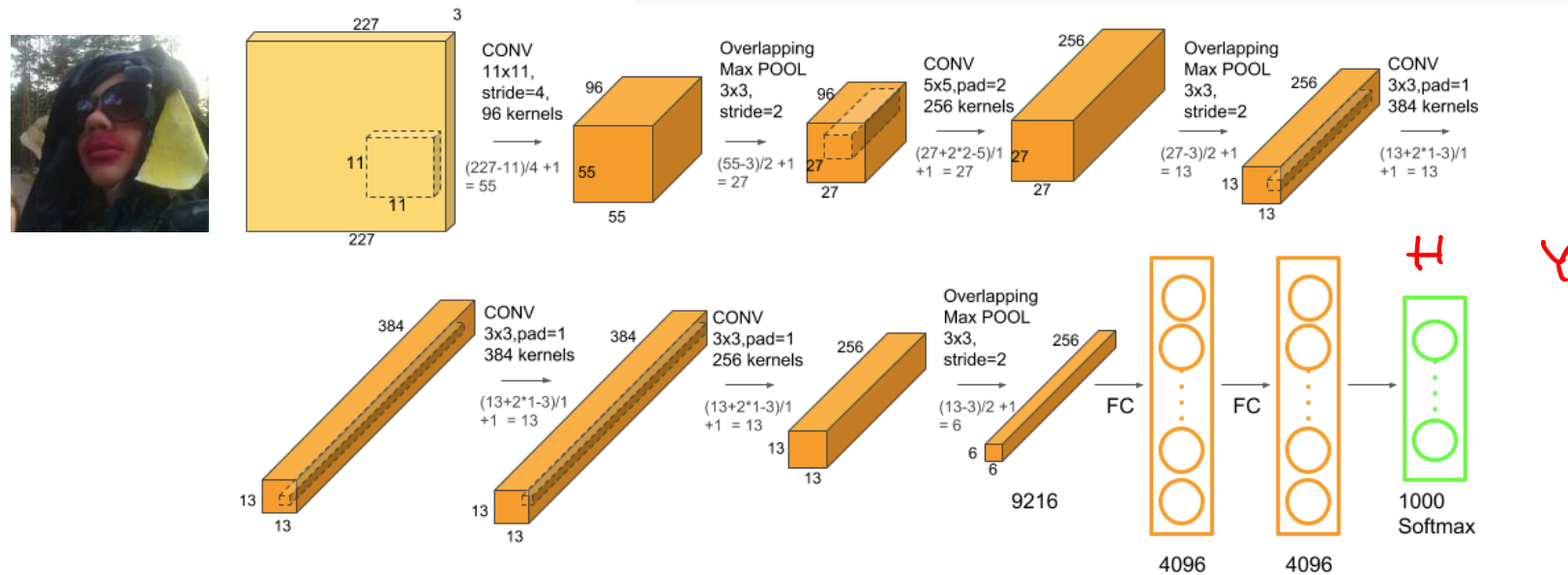
$$\Theta^{(k)} = \begin{pmatrix} \begin{pmatrix} \theta_{01}^{(k)} & \theta_{02}^{(k)} & \dots & \theta_{0n_{k+1}}^{(k)} \\ \theta_{11}^{(k)} & \theta_{12}^{(k)} & \dots & \theta_{1n_2}^{(k)} \\ \dots & \dots & \dots & \dots \\ \theta_{n_k1}^{(k)} & \theta_{n_k2}^{(k)} & \dots & \theta_{n_k n_{k+1}}^{(k)} \end{pmatrix} \end{pmatrix}.$$

Примечания / Notes:

- если на вход подаются большие изображения, то количество компонент матриц весов $\Theta^{(k)}$ становится велико;
- компоненты матриц весов $\Theta^{(k)}$ чувствительны к смещению объектов на изображении.

Классификация с помощью глубокой сверточной сети / Solving classification problem using deep convolutional neural network

Характерная архитектура сети. Размеры входного слоя соответствуют размерам изображения. Скрытые слои характеризуются значениями компонент матриц ~~фильтров~~ ^{ядер} (kernels). Несколько последних слоев полносвязные и характеризуются матрицами весов. На скрытых слоях, как правило, используется функция активации «relu». На выходном слое используется функция активации «softmax».

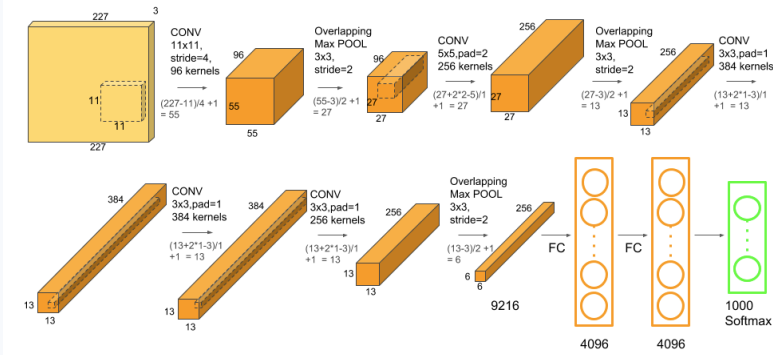


В процессе обучения минимизируется целевая функция ошибки с неизвестными значениями компонент матриц ядер $F^{(p)}$ слоев свертки и весов $\Theta^{(k)}$ полносвязных слоев сети:

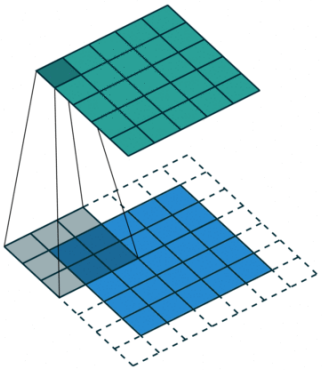
$$L(F^{(p)}, \Theta^{(k)}) = - \sum_{i=1}^m \sum_{j=1}^{n_l=1000} \left(y_j^{(i)} \ln \left(h_j^{(i)} \right) \right) \Rightarrow \min.$$

Классификация с помощью глубокой сверточной сети / Solving classification problem using deep convolutional neural network

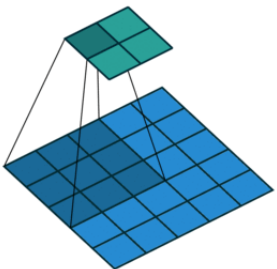
Характерная архитектура сети. Размеры входного слоя соответствуют размерам изображения. Скрытые слои характеризуются значениями компонент матриц ядер (**kernels**). Несколько последних слоев полносвязные и характеризуются матрицами весов. На скрытых слоях, как правило, используется функция активации «**relu**». На выходном слое используется функция активации «**softmax**».



Дополнение / **Padding**



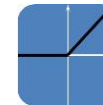
Шагание / **Striding**



Свертка / **Convolution**

3 ₀	3 ₁	2 ₂	1	0
0 ₂	0 ₂	1 ₀	3	1
3 ₀	1 ₁	2 ₂	2	3
2	0	0	2	2
2	0	0	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0



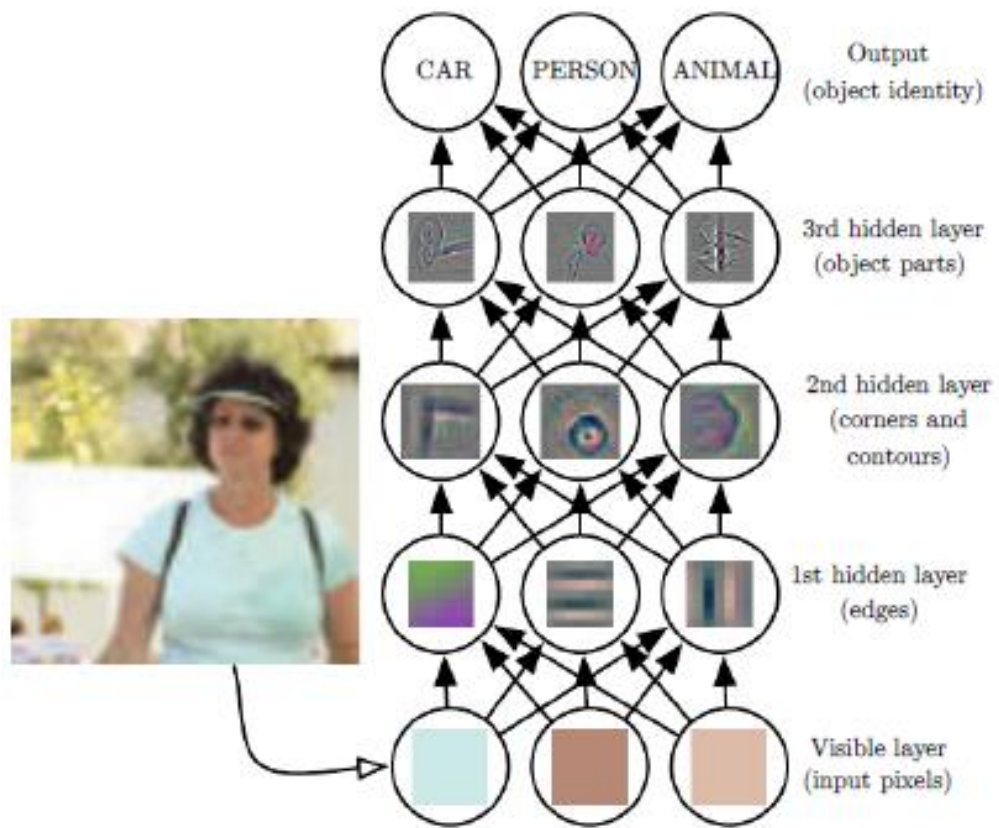
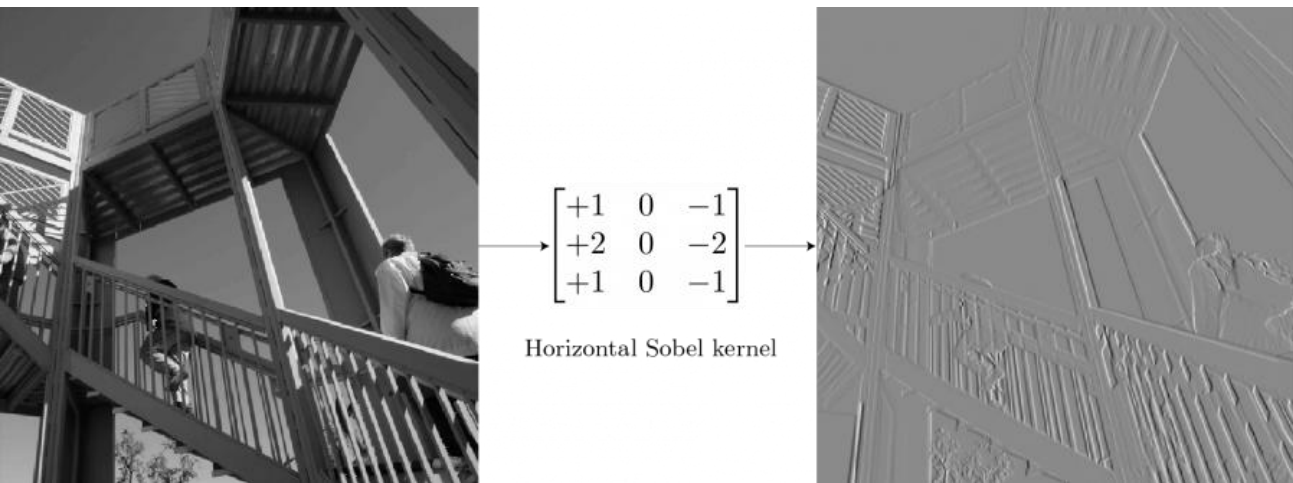
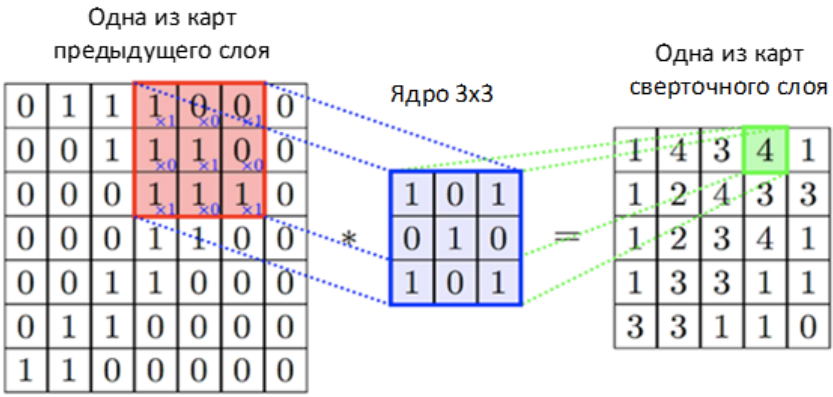
0 1 2
2 2 0
0 1 2

Группирование / **Pooling**

Max pool
filter [2,2]
stride [1,1]

17	19
17	19

Свертка / Convolution

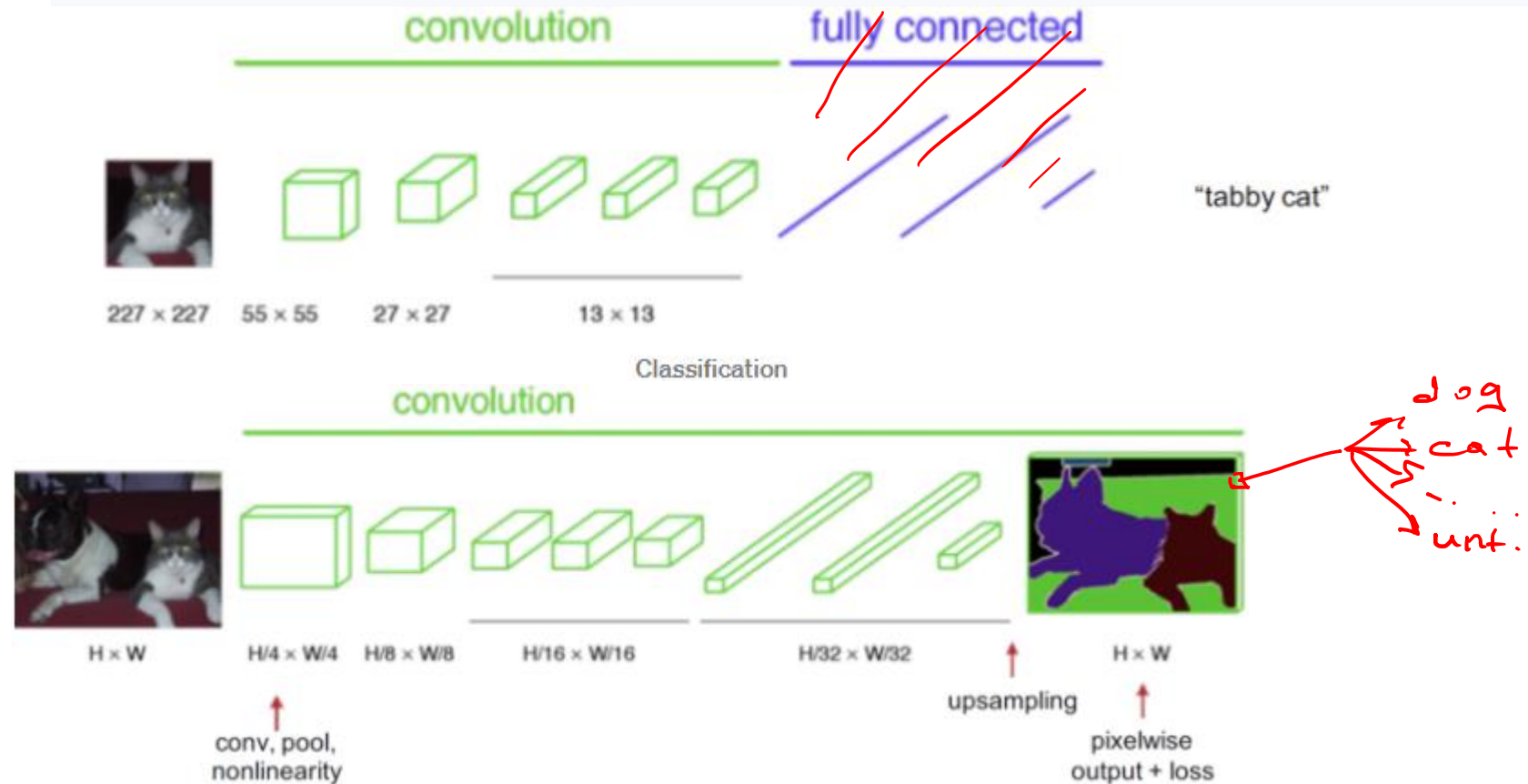


Сегментация как случай классификации / Semantic segmentation is a particular case of a classification problem

Характерная архитектура сети. Размеры входного слоя соответствуют размерам изображения. Скрытые слои характеризуются значениями компонент матриц ядер (**kernels**). Полносвязные слои, как правило, отсутствуют. На скрытых слоях, как правило, используется функция активации «**relu**». На выходном слое используется функция активации «**softmax**».

Примечания / Notes:

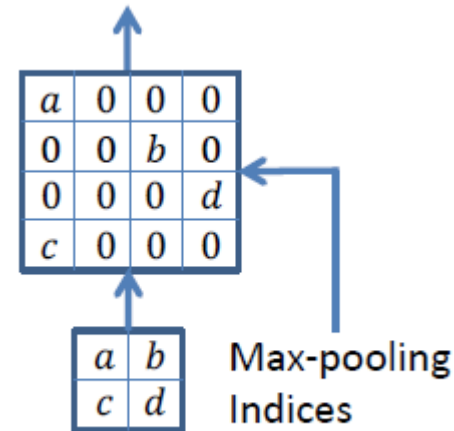
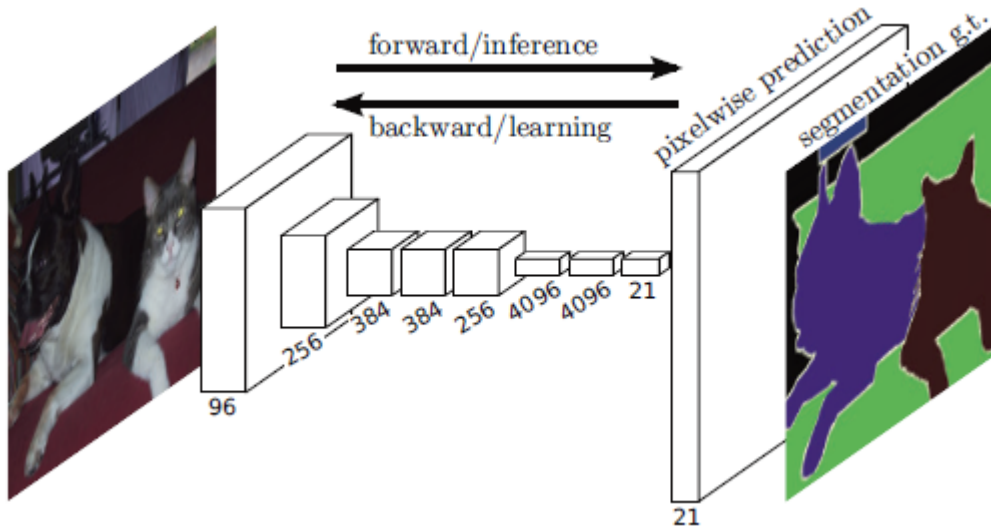
- семантическая сегментация не требует больших датасетов;
- обученная сеть может сегментировать изображения не только равного, но и большего разрешения;
- все это благодаря отсутствию полносвязных слоев в сети;
- семантическая сегментация требует разметки изображений, часто ручную.



$$L(F^{(p)}) = - \sum_{i=1}^m \sum_{j=1}^{n_l} \left(y_j^{(i)} \ln \left(h_j^{(i)} \right) \right) \Rightarrow \min.$$

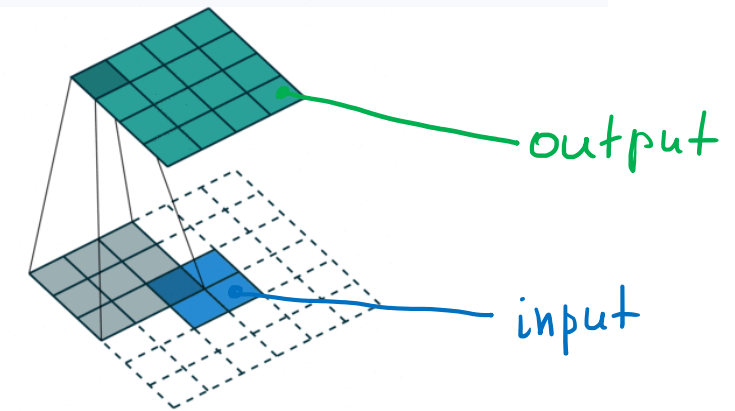
Сегментация как случай классификации / Semantic segmentation is a particular case of a classification problem

Семантическая сегментация. Размеры входного слоя соответствуют размерам изображения. Скрытые слои характеризуются значениями компонент матриц ядер (**kernels**). Полносвязные слои, как правило, отсутствуют. На скрытых слоях, как правило, используется функция активации «**relu**». На выходном слое используется функция активации «**softmax**».



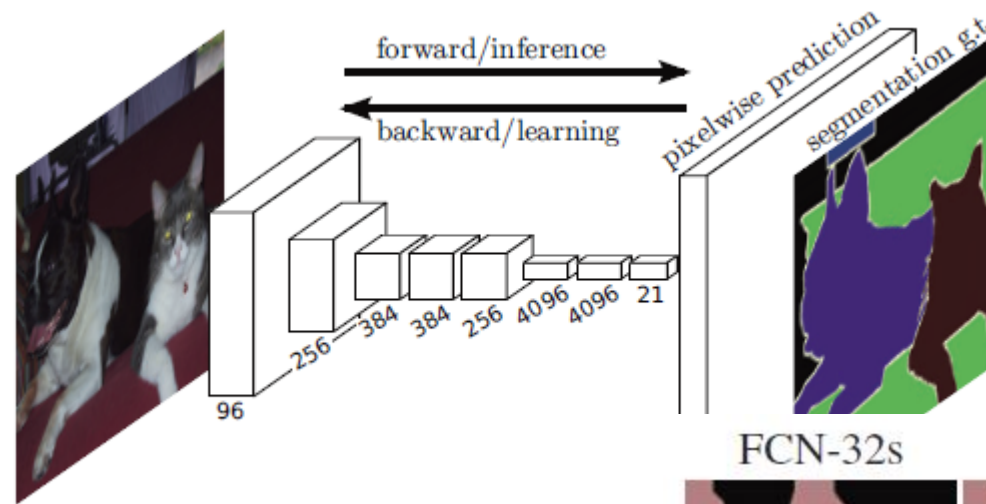
Upsampling Using Max-Pooling Indices

Обратная свертка / Transposed convolution



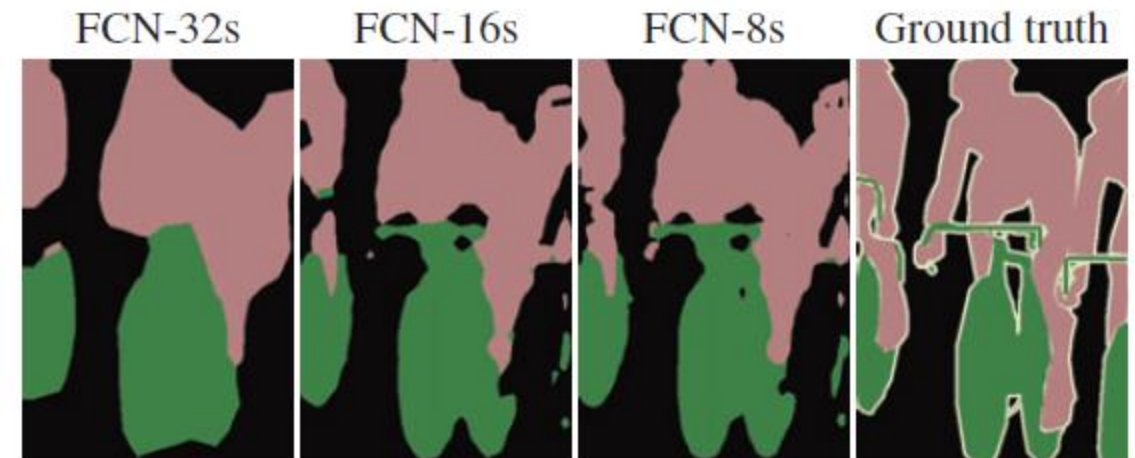
Сегментация как случай классификации / Semantic segmentation is a particular case of a classification problem

Семантическая сегментация. Размеры входного слоя соответствуют размерам изображения. Скрытые слои характеризуются значениями компонент матриц ядер (**kernels**). Полносвязные слои, как правило, отсутствуют. На скрытых слоях, как правило, используется функция активации «**relu**». На выходном слое используется функция активации «**softmax**».



$$\text{Intersection over Union (IOU) or Jaccard Index} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

The diagram shows two overlapping blue squares. The top square represents the 'Area of Overlap' (Intersection), and the bottom square represents the 'Area of Union' (the combined area of both squares). The formula calculates the ratio of the intersection area to the union area.

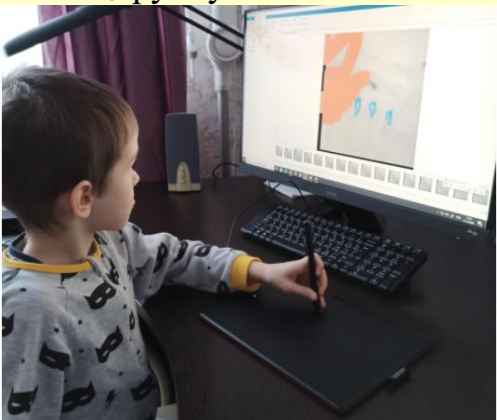


Comparison with different FCNs

[Review on FCN](#)

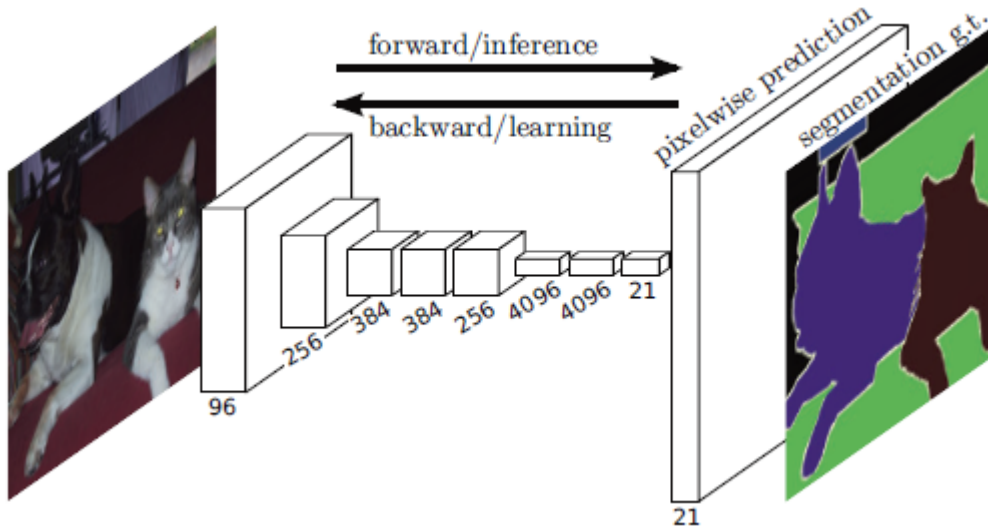
Примечания / Notes:

- семантическая сегментация не требует больших датасетов;
- обученная сеть может сегментировать изображения не только равного, но и большего разрешения;
- все это благодаря отсутствию полносвязных слоев в сети;
- семантическая сегментация требует разметки изображений, часто вручную.

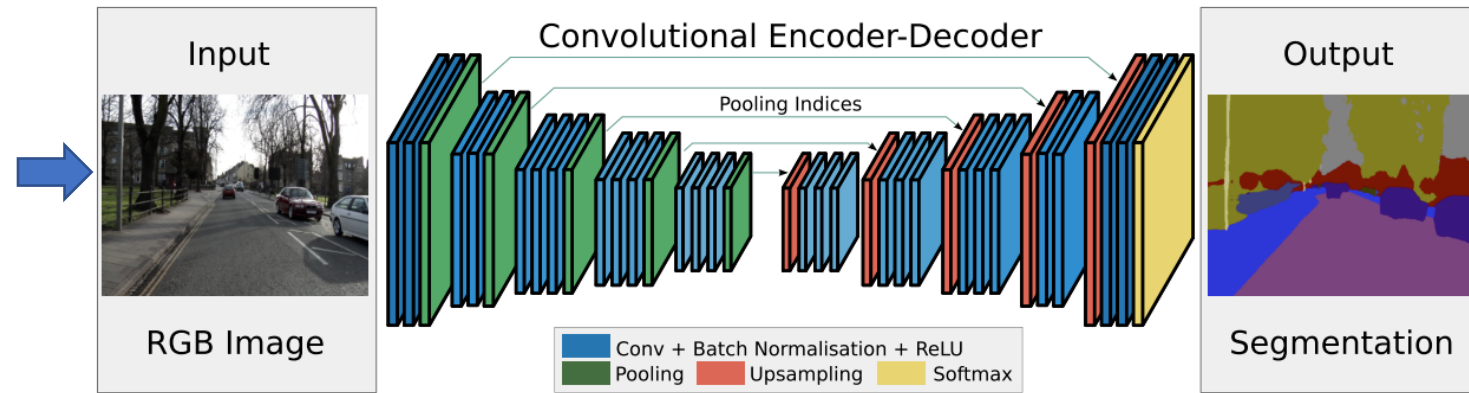


Сегментация как случай классификации / Semantic segmentation is a particular case of a classification problem

Семантическая сегментация. Размеры входного слоя соответствуют размерам изображения. Скрытые слои характеризуются значениями компонент матриц ядер (**kernels**). Полносвязные слои, как правило, отсутствуют. На скрытых слоях, как правило, используется функция активации «**relu**». На выходном слое используется функция активации «**softmax**».



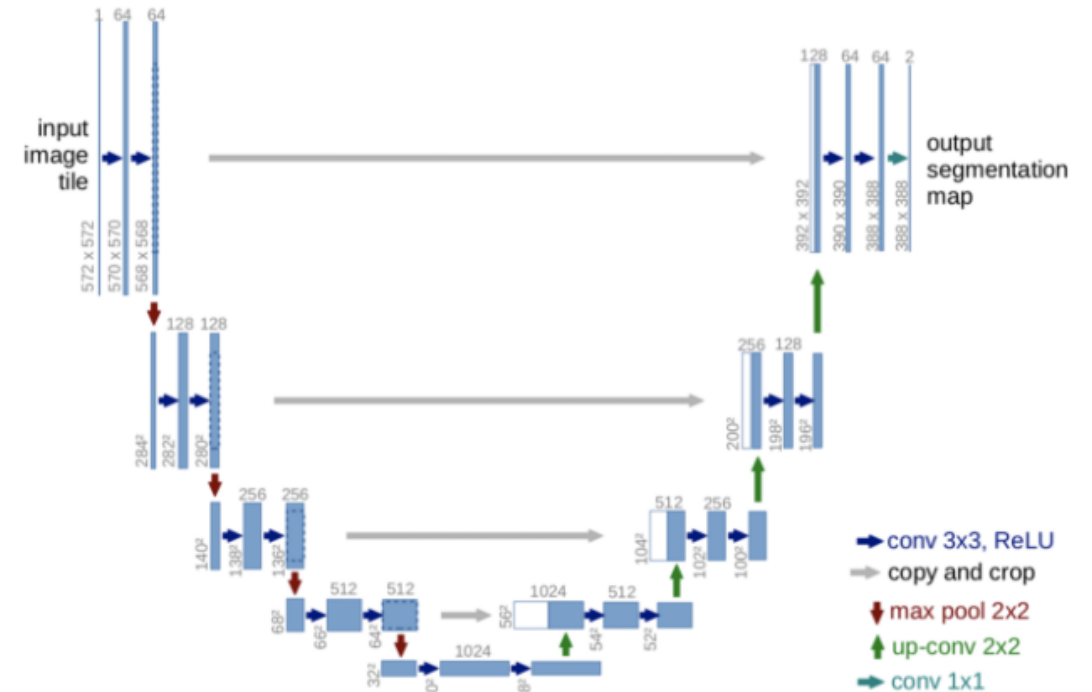
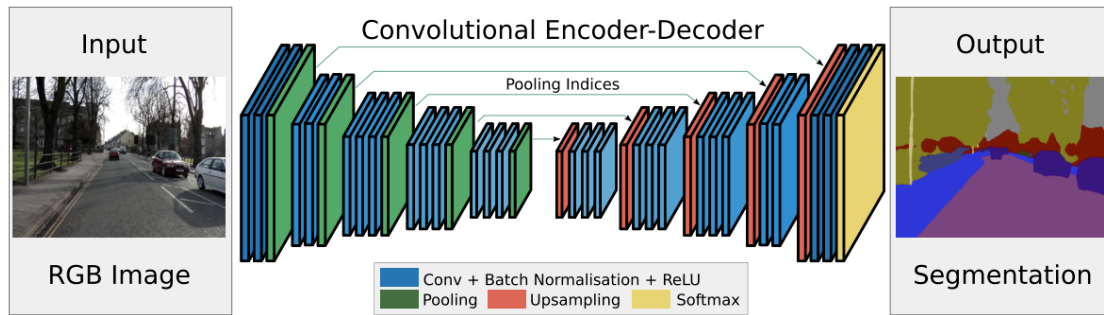
[Review on FCN](#)



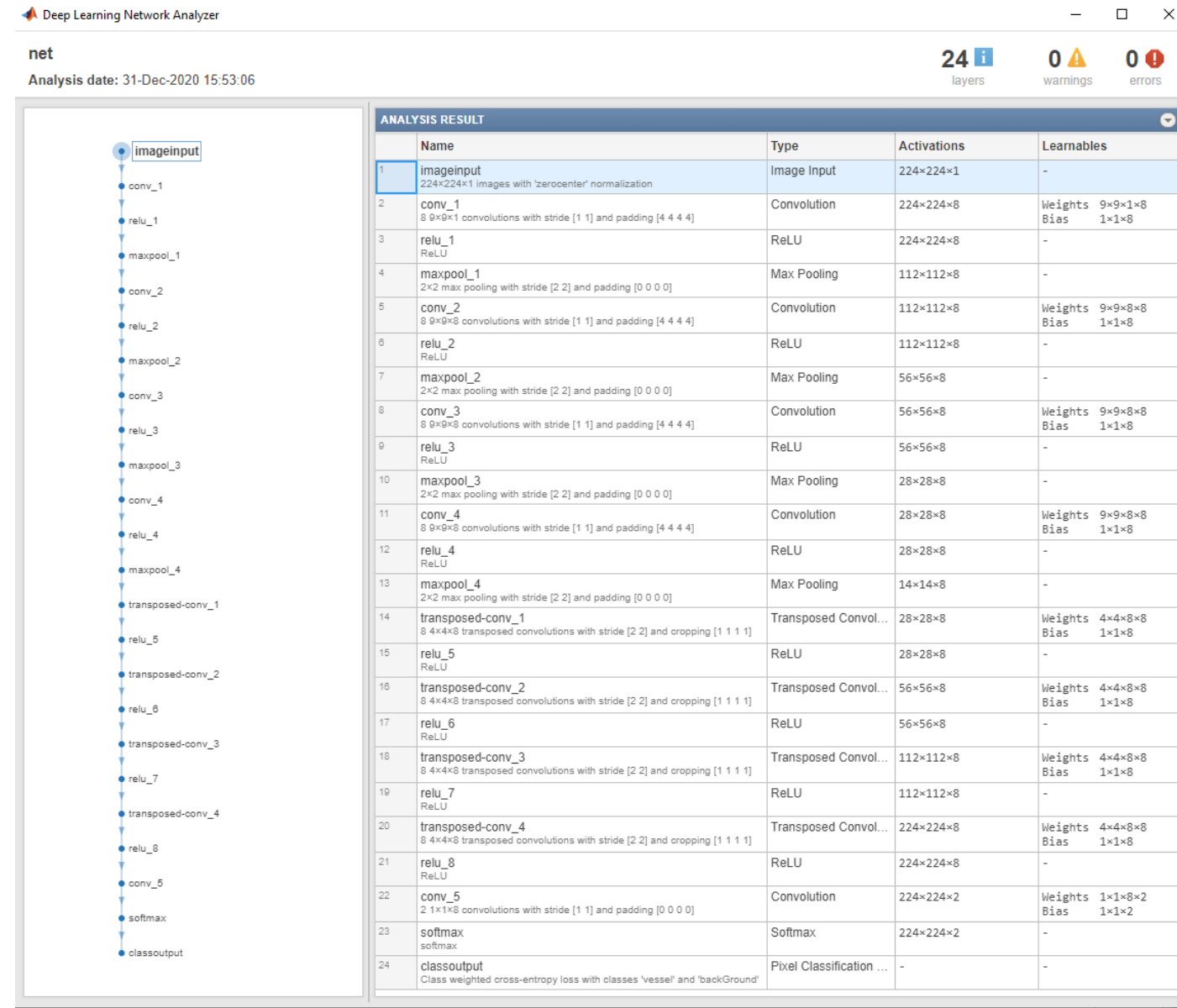
[Review on SegNet](#)

Сегментация как случай классификации / Semantic segmentation is a particular case of a classification problem

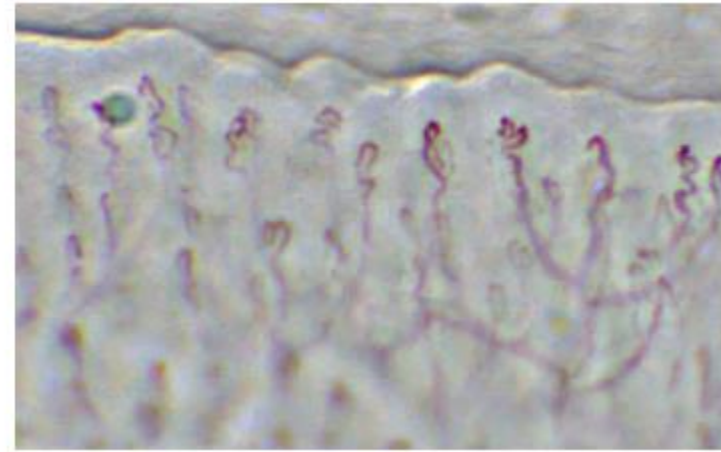
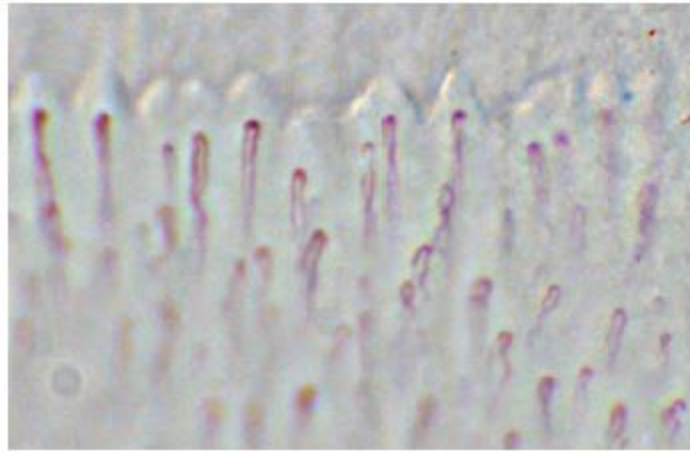
Семантическая сегментация. Размеры входного слоя соответствуют размерам изображения. Скрытые слои характеризуются значениями компонент матриц ядер (**kernels**). Полносвязные слои, как правило, отсутствуют. На скрытых слоях, как правило, используется функция активации «**relu**». На выходном слое используется функция активации «**softmax**».



Разработка сети в MATLAB / Semantic segmentation using MATLAB



Разработка сети в MATLAB / Semantic segmentation using MATLAB



Workspace	
Name ▲	Value

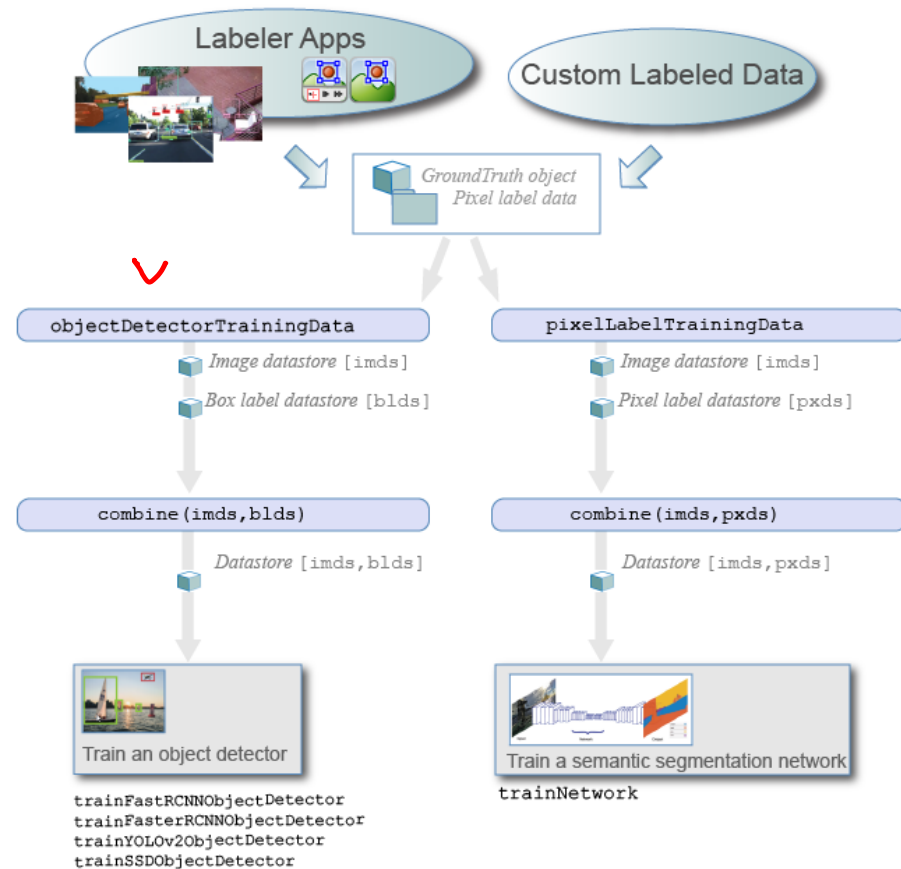
```
1 clear
2 clc
3
4 load res_221
5
6 imDir='F:\st
7
8 cd %'F:\stud
9 saveDir=cd;
```



Самостоятельная работа / Homework

Вопросы и задания.

1. Можно ли использовать предобученные сети в решении задач семантической сегментации?
2. Основы Object detection.



Полезные ссылки / **Links**

Онлайн курсы, обучающие ресурсы:

[Deep Learning на пальцах](#): видео курс, плюс материалы на русском языке от С.Козлова