



# **Свёрточные нейронные сети.**

Евгений Борисов

# Свёрточные нейросети

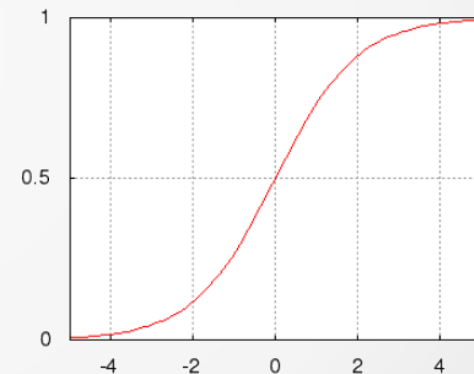
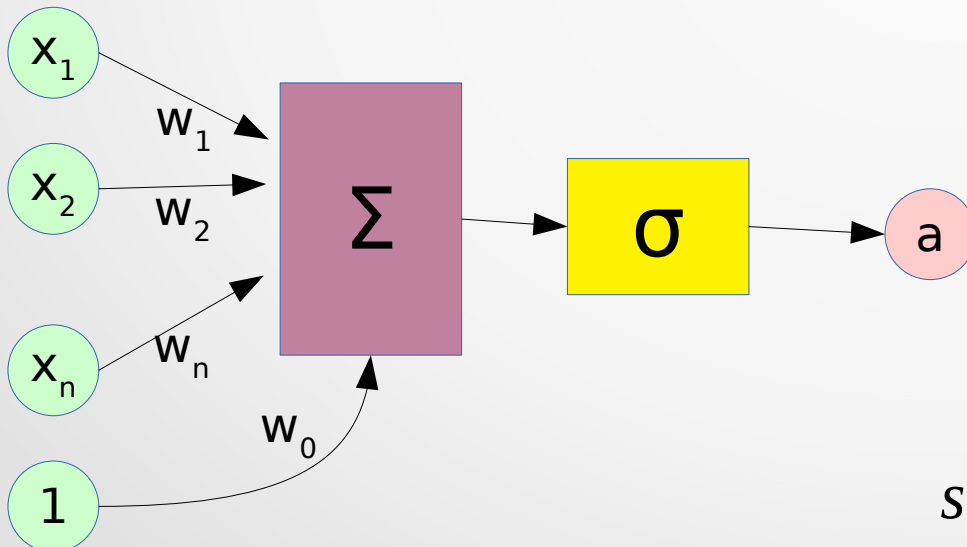
## модель нейрона

$$a(x, w) = \sigma\left(\sum_{i=1}^n x_i \cdot w_i - w_0\right) = \sigma(\langle x, w \rangle)$$

$x_i$  - ВХОД

$w_i$  - ВЕС СВЯЗИ

$\sigma$  - функция активации



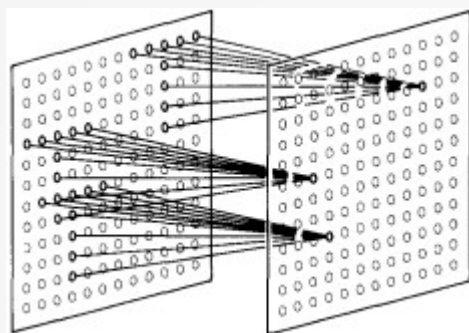
состояние нейрона

$$s(x, w) = \sum_{i=1}^n x_i \cdot w_i - w_0$$

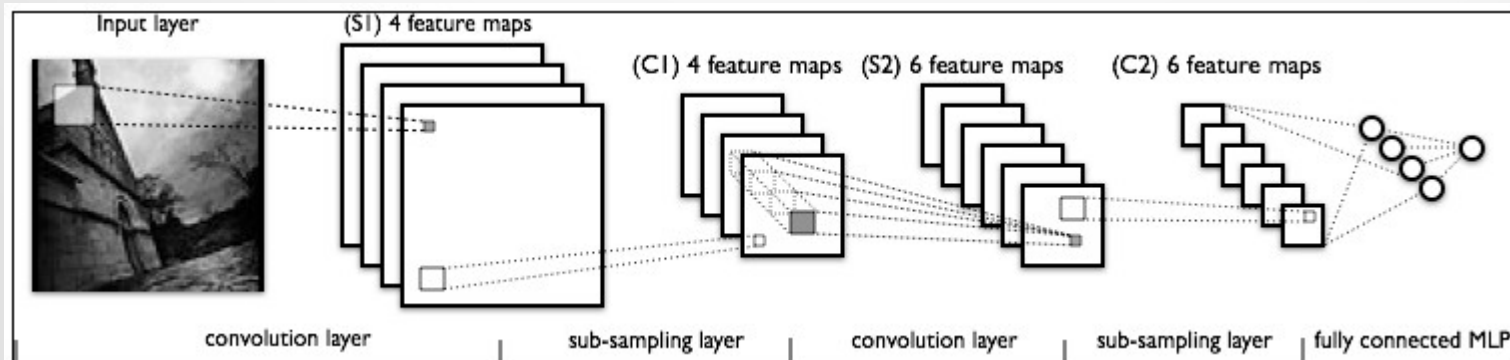
# Свёрточные нейросети

## Свёрточные сети

Fukushima, Neocognitron (1980). "A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position". Biological Cybernetics. 36 (4): 193–202. doi:10.1007/bf00344251.



Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel: Backpropagation Applied to Handwritten Zip Code Recognition, Neural Computation, 1(4):541-551, Winter 1989.



# Свёрточные нейросети

## Операция свёртки

$$(f * g)[m, n] = \sum_{k, l} f[m - k, n - l] \cdot g[k, l]$$

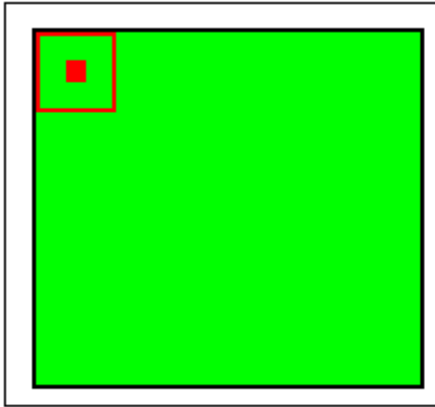


Рис.2: обработка краёв valid

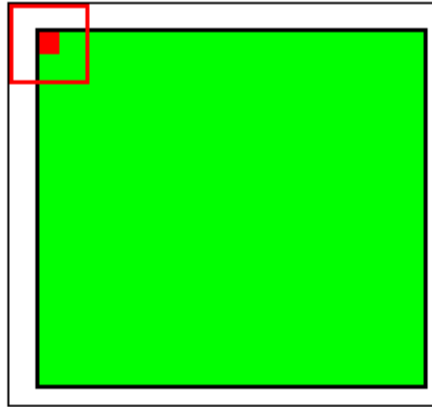


Рис.3: обработка краёв same

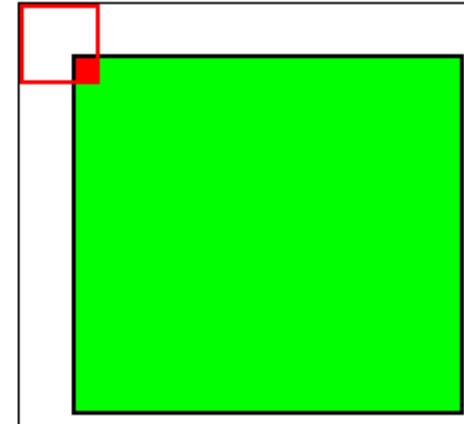


Рис.4: обработка краёв full

g -ядро свёртки

- берём точку с окрестностью,
- поэлементно умножаем эту матрицу на ядро, результат суммируется и записывается как новое значение данной точки
- процедура повторяется для всех точек изображения.

# Свёрточные нейросети

## примеры ядер свёртки

$$g = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ копирование (без изменений)}$$

$$g = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \text{ сдвиг влево на 1 пиксел}$$

$$g = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \text{ сглаживание или усреднение по окрестности (box filter)}$$

# Свёрточные нейросети

## примеры ядер свёртки



Рис.11: исходная картинка



Рис.12: сглаженная картинка

сглаживание или усреднение  
по окрестности (box filter)



результат применения  
фильтра Собеля

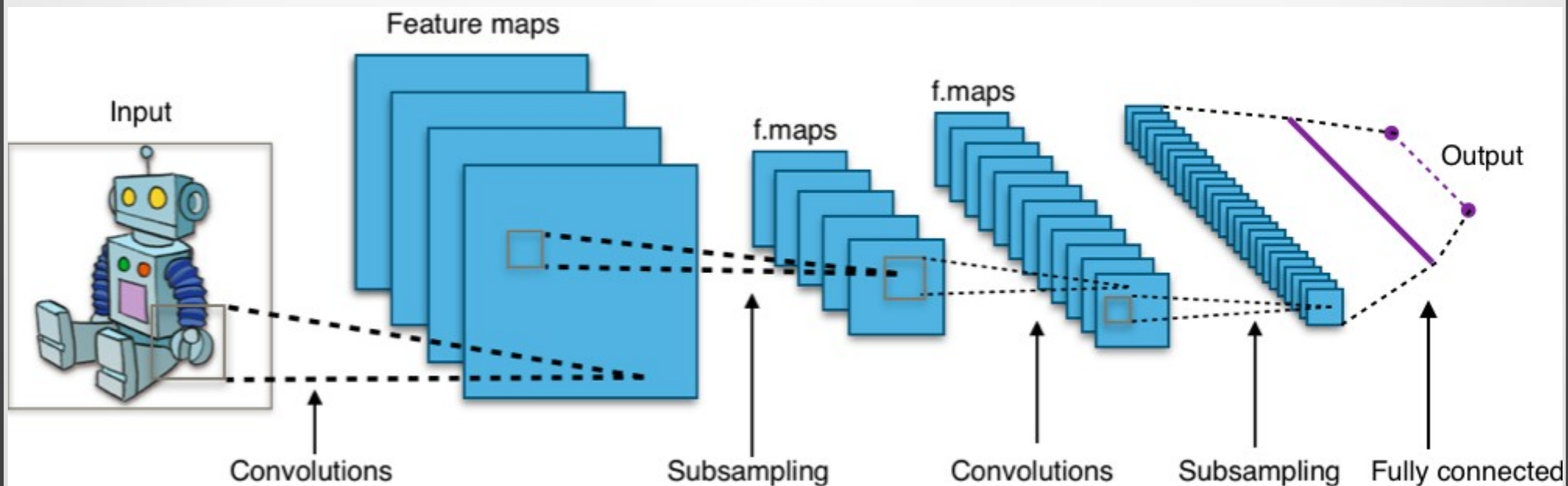
# Свёрточные нейросети

## Свёрточная сеть

свёрточный слой (convolution)

слой подвыборки (subsampling)

слой MLP

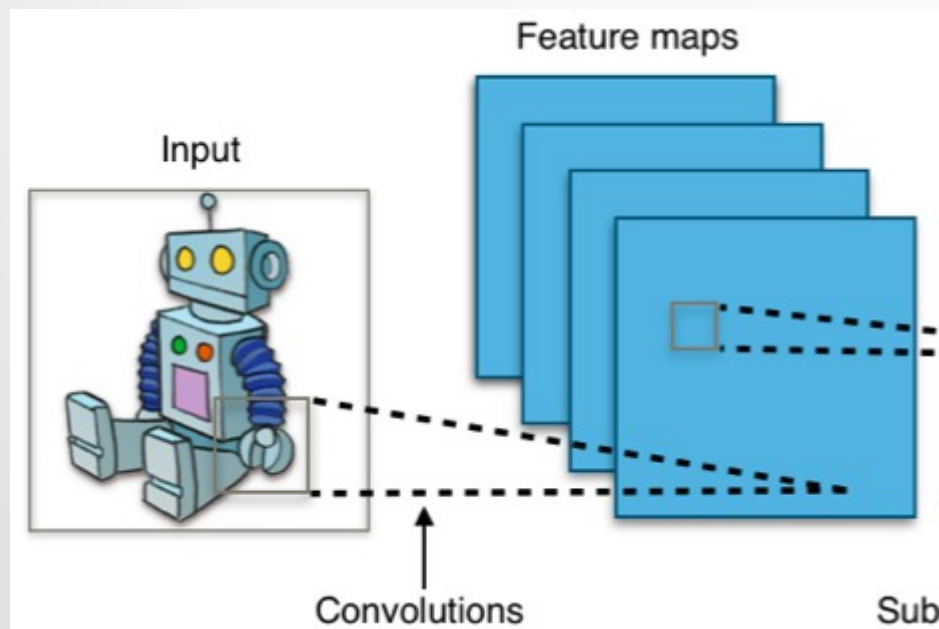


# Свёрточные нейросети

## Свёрточный слой

$$x_j^l = f \left( \sum_i x_i^{l-1} * k_j^l + b_j^l \right)$$

n ядер  
выполняем свёртку  
получаем n карт признаков





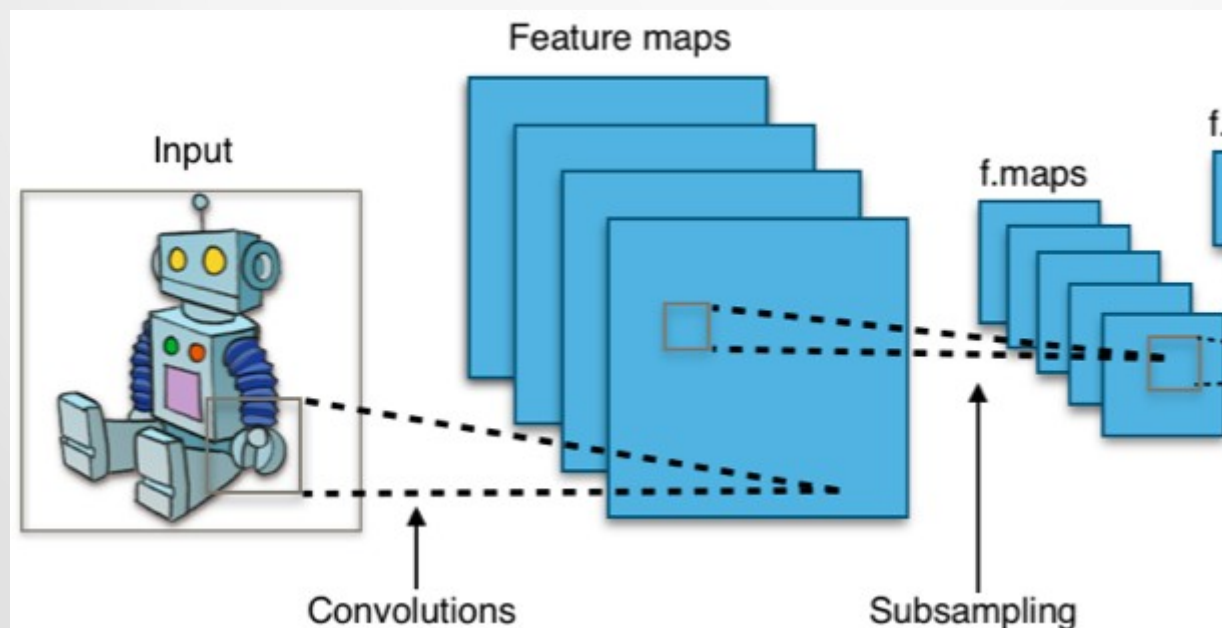
# Свёрточные нейросети

## слой подвыборки (subsampling)

уменьшение размера входной карты признаков (обычно в 2 раза).  
методом выбора максимального элемента (max-pooling)

карта признаков разделяется на ячейки 2x2 элемента  
из ячеек выбираем максимальные по значению

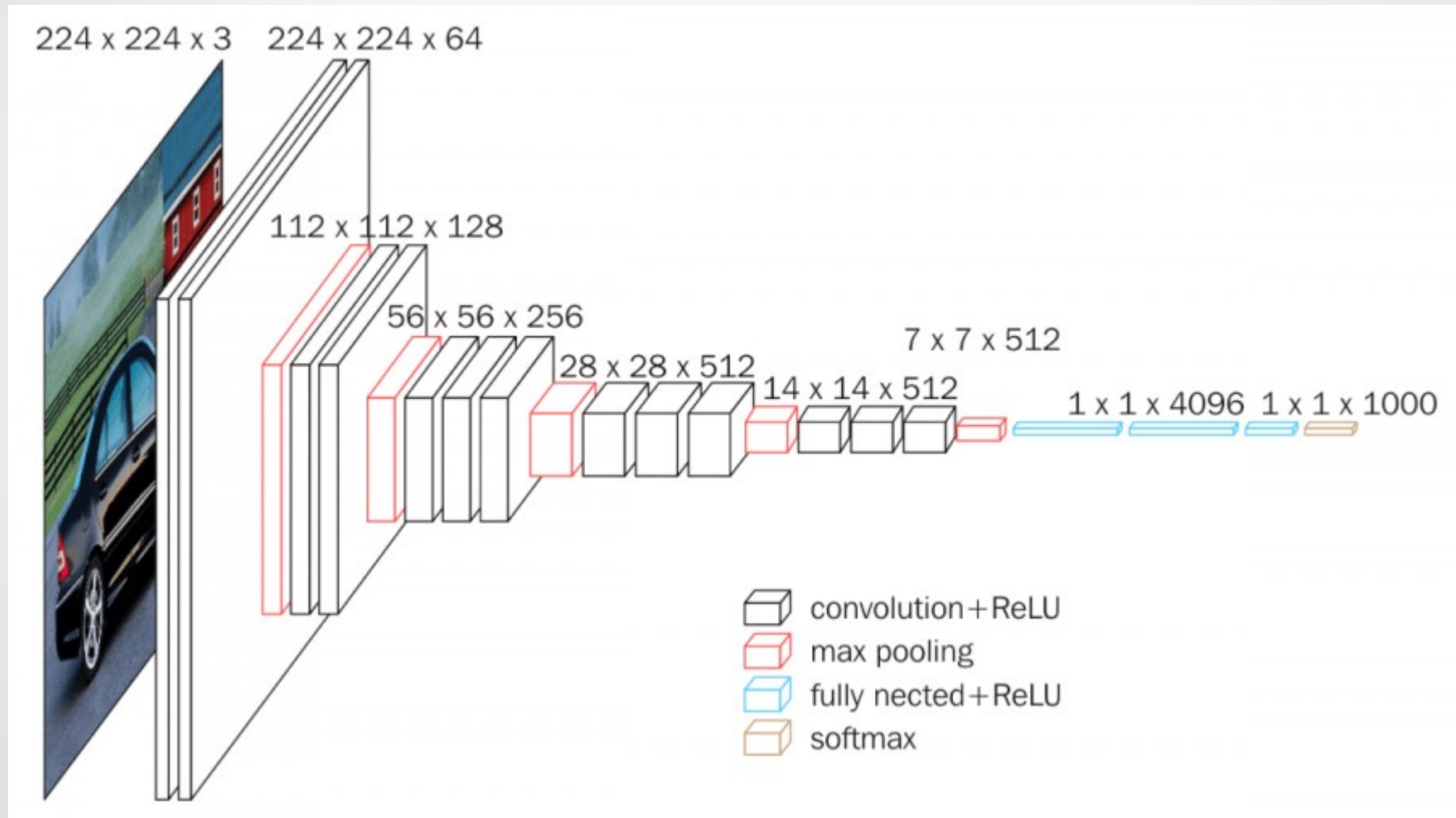
$$x^l = f(a^l \cdot \text{subsample}(x^{l-1}) + b^l)$$



# Свёрточные нейросети

## VGG-16

<https://arxiv.org/pdf/1409.1556.pdf>



# Свёрточные нейросети

## ResNet (Residual Network)

<https://arxiv.org/pdf/1512.03385.pdf>

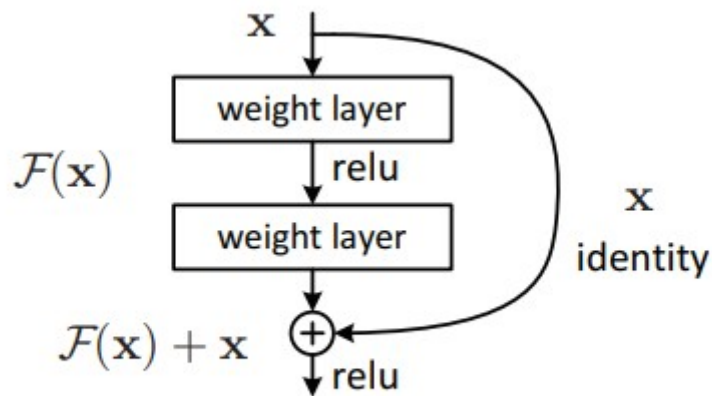
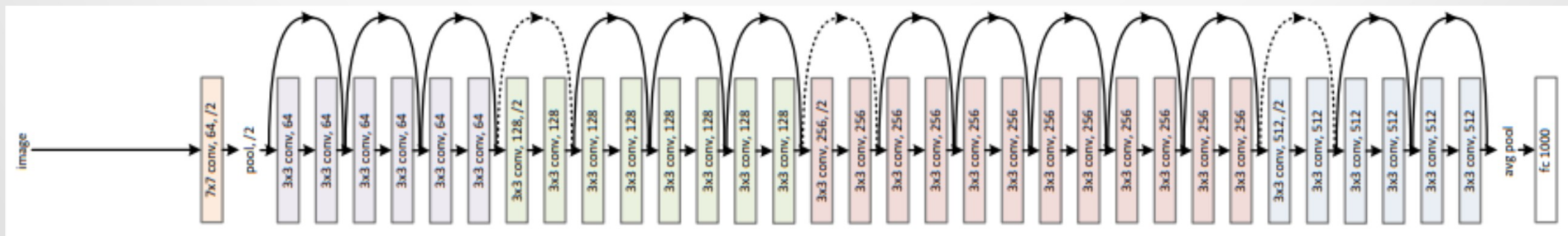


Figure 2. Residual learning: a building block.

соединения быстрого доступа



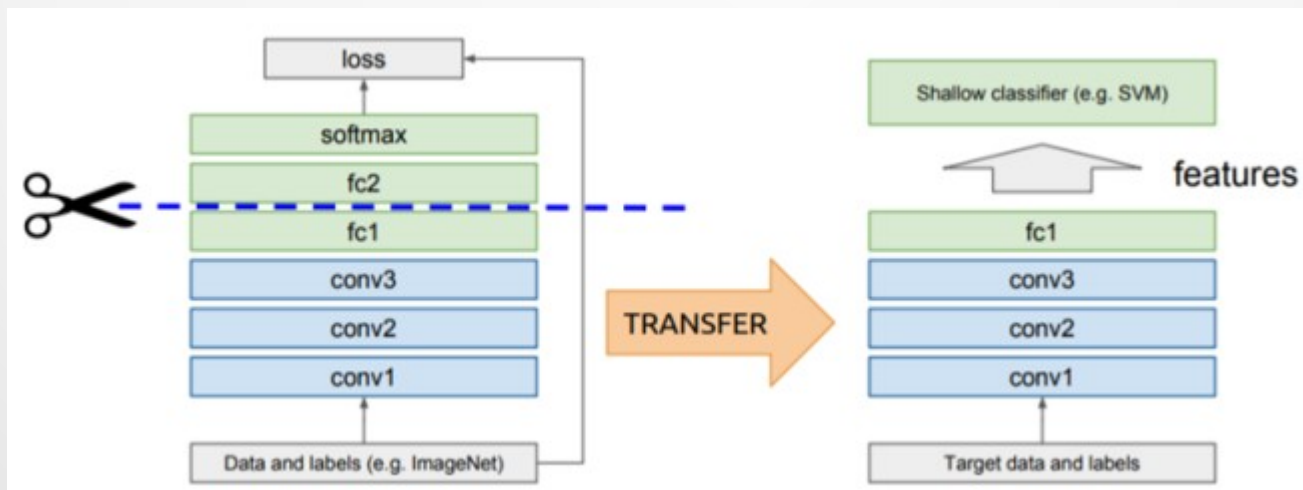
# Свёрточные нейросети

**Transfer learning** — перенос частей обученных ИНС в другие модели

Jason Yosinski, Je Clune, Yoshua Bengio, Hod Lipson. How transferable are features in deep neural networks? 2014.

Свёрточная сеть для обработки изображений:

- $z = f(x, \alpha)$  — свёрточные слои для векторизации объектов
- $y = g(z, \beta)$  — полносвязные слои под конкретную задачу



# Поиск объектов на изображении

## Модель фона с помощью нейросети

семантическая сегментация



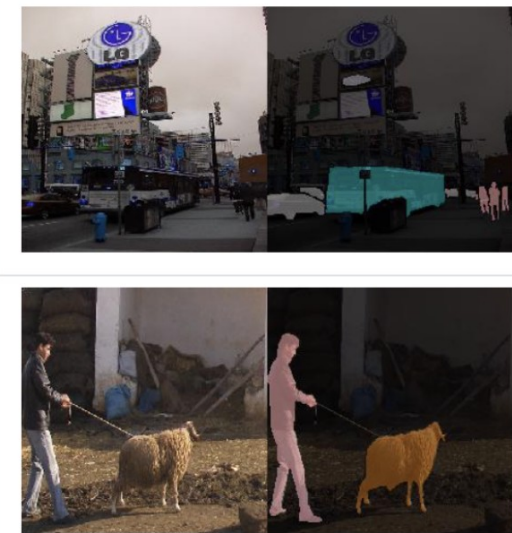
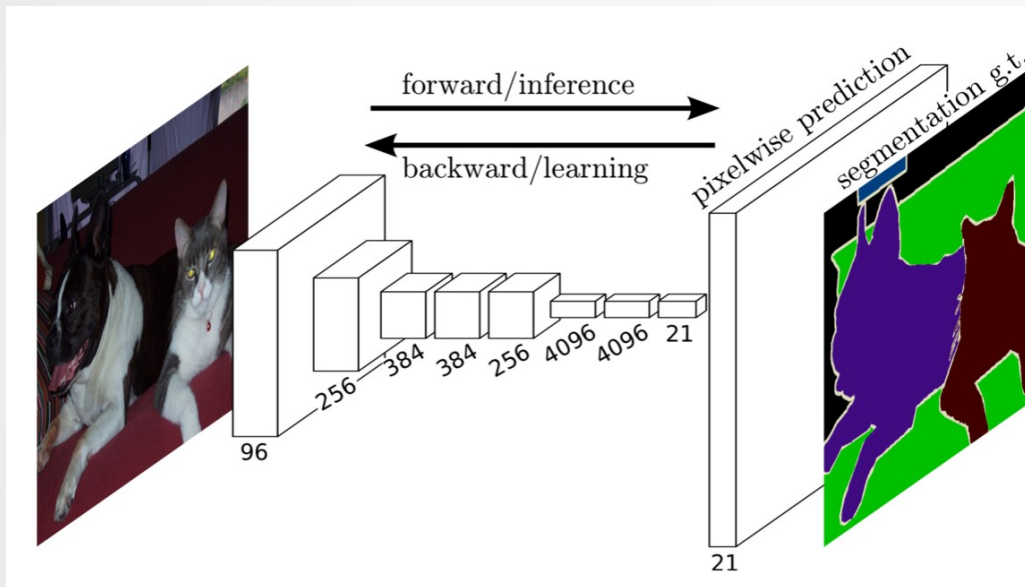


# Поиск объектов на изображении

## Semantic segmentation

FCN: Fully Convolutional Networks

<https://arxiv.org/pdf/1411.4038.pdf>



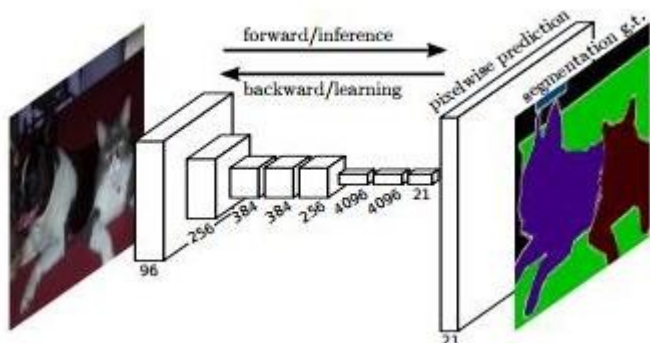
выход — карты поточечной оценки

для каждого класса своя карта

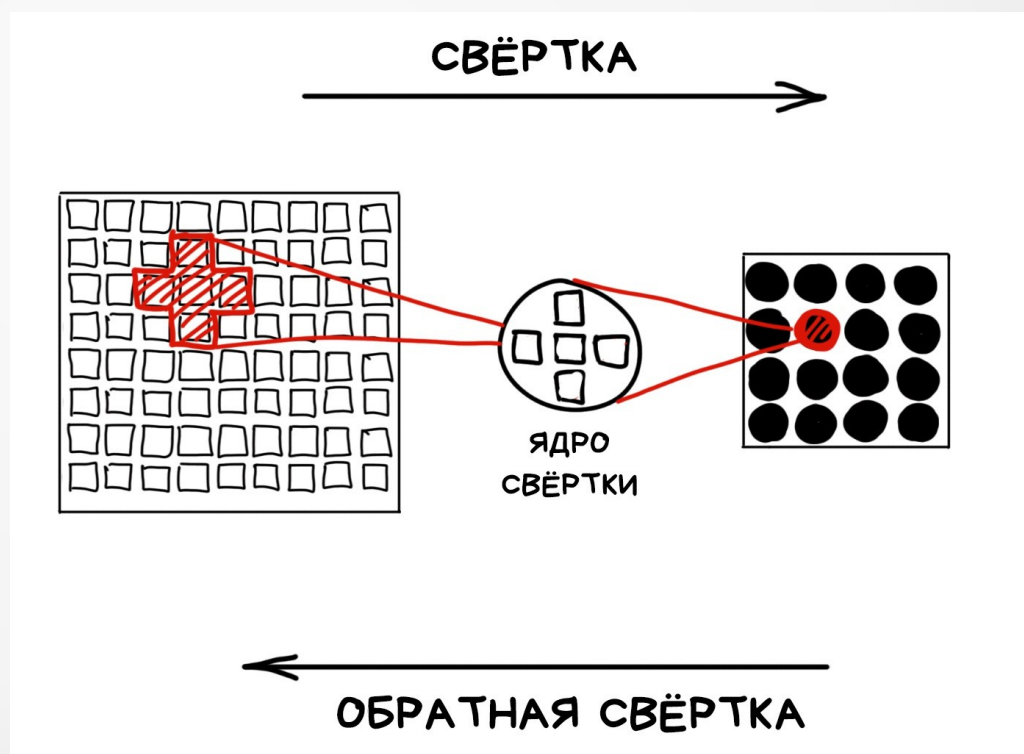
размер входного изображения = размеру входной карты

сравниваем выходные карты поточечно,  
для каждой точки определяем карту-победителя

# Поиск объектов на изображении



изображение обрабатывается свёрточными слоями  
на выходе выполняем обратную свёртку



# Поиск объектов на изображениях

## FCN: Fully Convolutional Networks

Пример — ищем людей на картинке (датасет Pascal VOC)

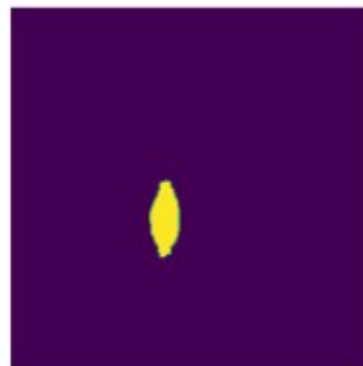
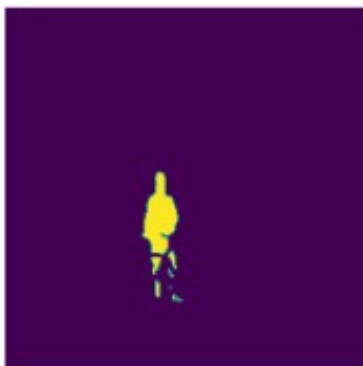
картинка



разметка



результат FCN

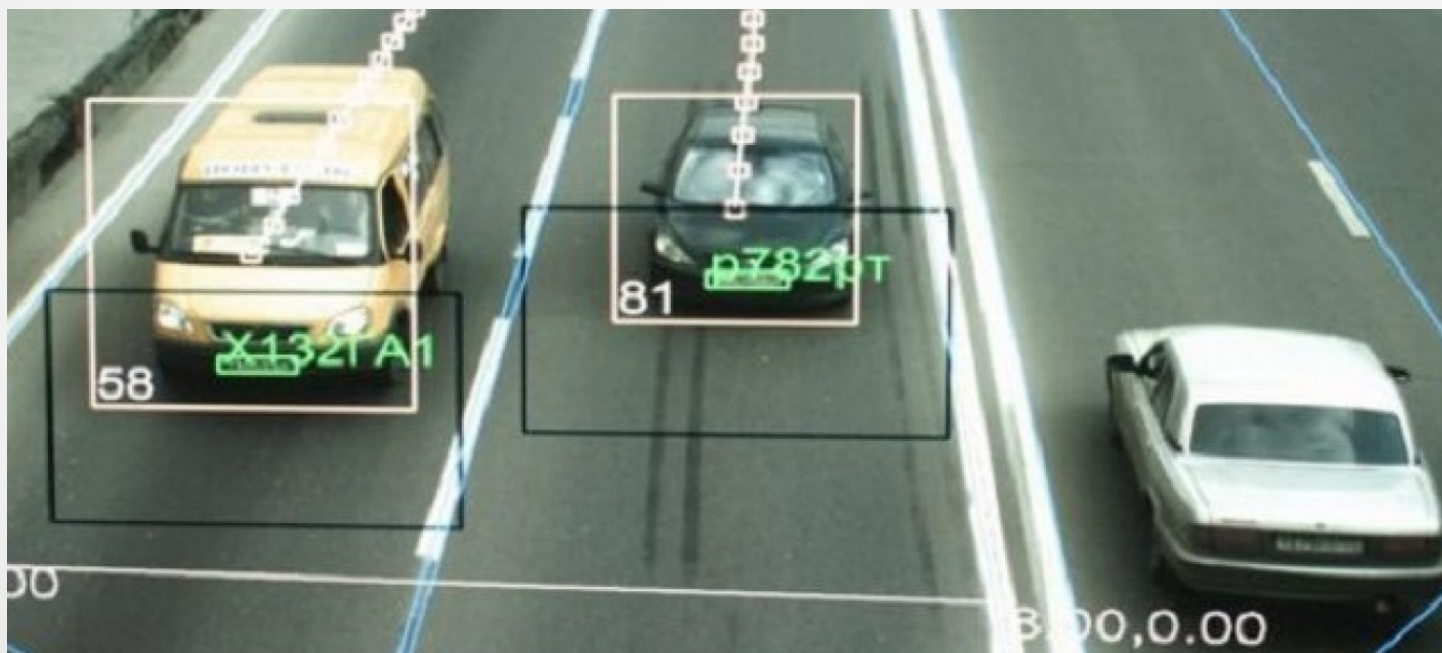




# Поиск объектов на изображении

## Модель объекта с помощью нейросети

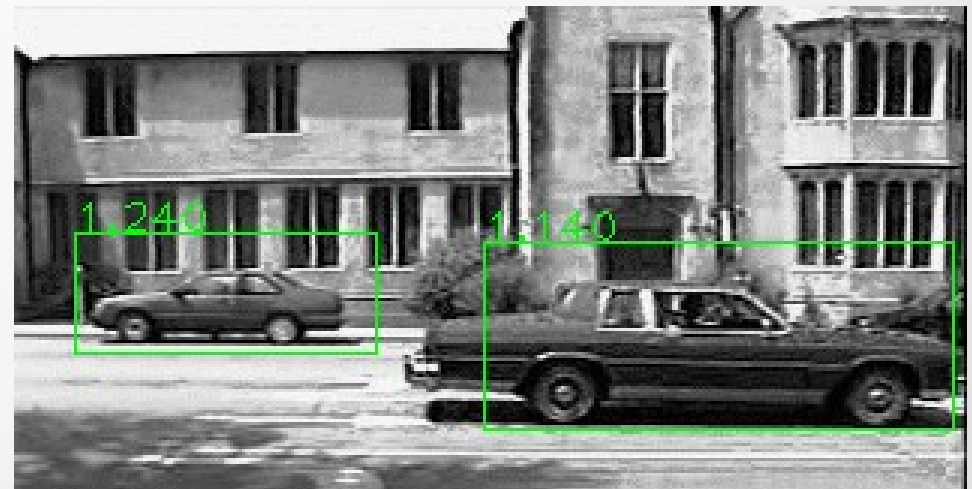
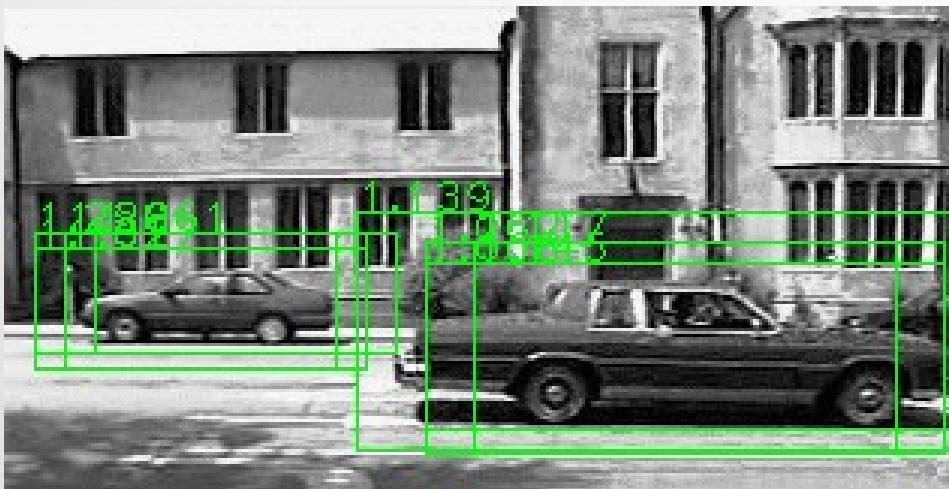
Задача классификации



# Поиск объектов на изображении

## Локализация объектов : метод скользящего окна

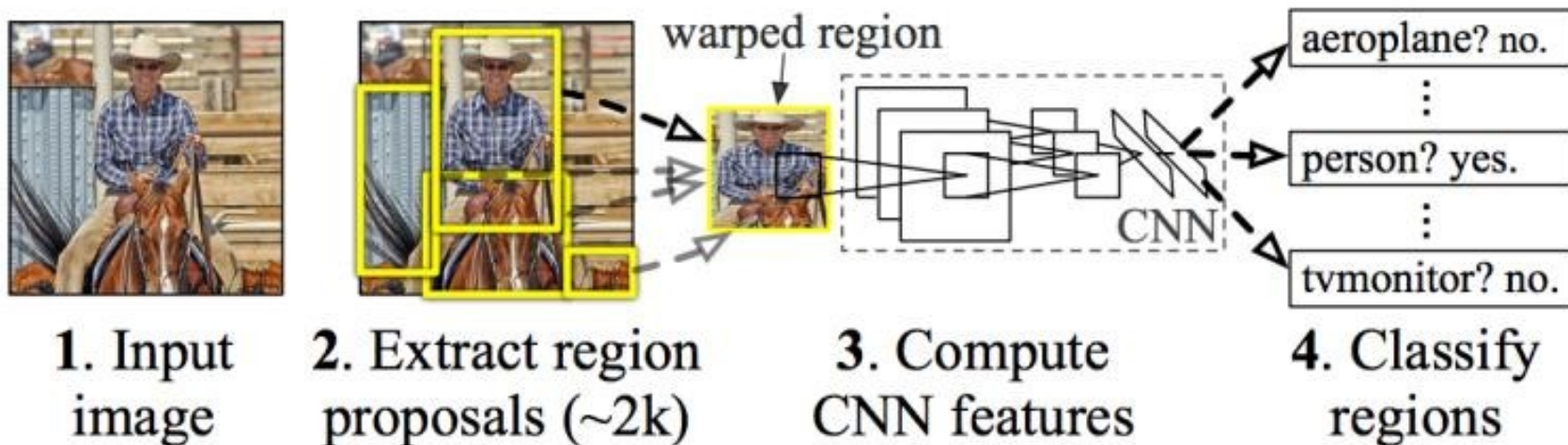
1. задать размер окна
2. пройти окном изображение
3. на каждом шаге выполняем классификацию содержимого окна
4. изменить размер окна и повторить процедуру с п.2
5. обрабатываем результаты



# Поиск объектов на изображениях

## Object detection

### R-CNN: *Regions with CNN features*



### Region Based Convolutional Neural Networks (R-CNN)

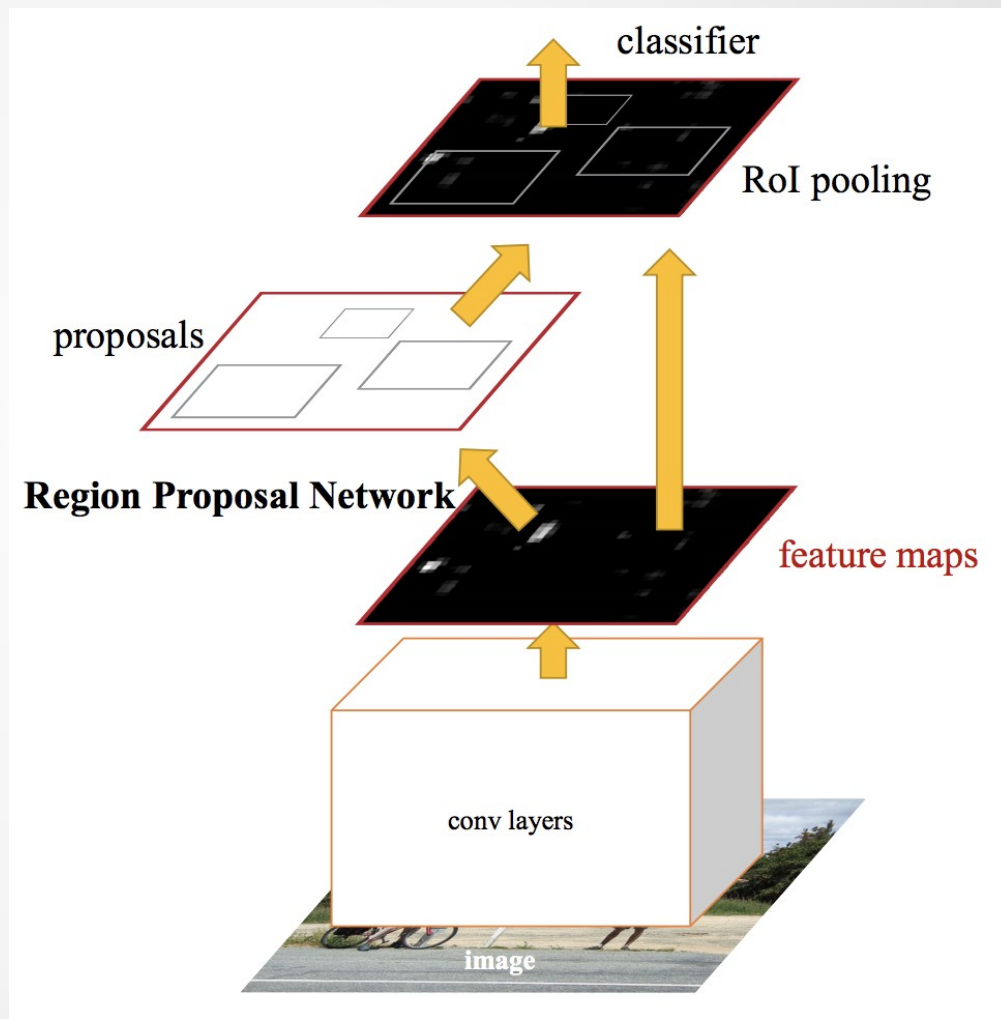
изображение разделяется на части

каждую часть проверяем классификатором

# Поиск объектов на изображении

## Faster-R-CNN

- принимаем картинку на вход
- картинка прогоняется через CNN, формируем feature maps
- определяем регионы-кандидаты (возможно содержащие объекты)
- выделяем эти регионы
- и применяем к ним классификатор картинок



# Свёрточные нейросети

## **Датасеты**

MNIST (National Institute of Standards and Technology)

CIFAR-10 (Canadian Institute for Advanced Research)

Pascal VOC (Visual Object Classes )

ImageNet

# Свёрточные нейросети

## MNIST (National Institute of Standards and Technology)

28x28 grayscale, 60K training images, 10K testing images, 10 classes

<https://github.com/zalandoresearch/fashion-mnist>

[https://ru.wikipedia.org/wiki/MNIST\\_\(база\\_данных\)](https://ru.wikipedia.org/wiki/MNIST_(база_данных))

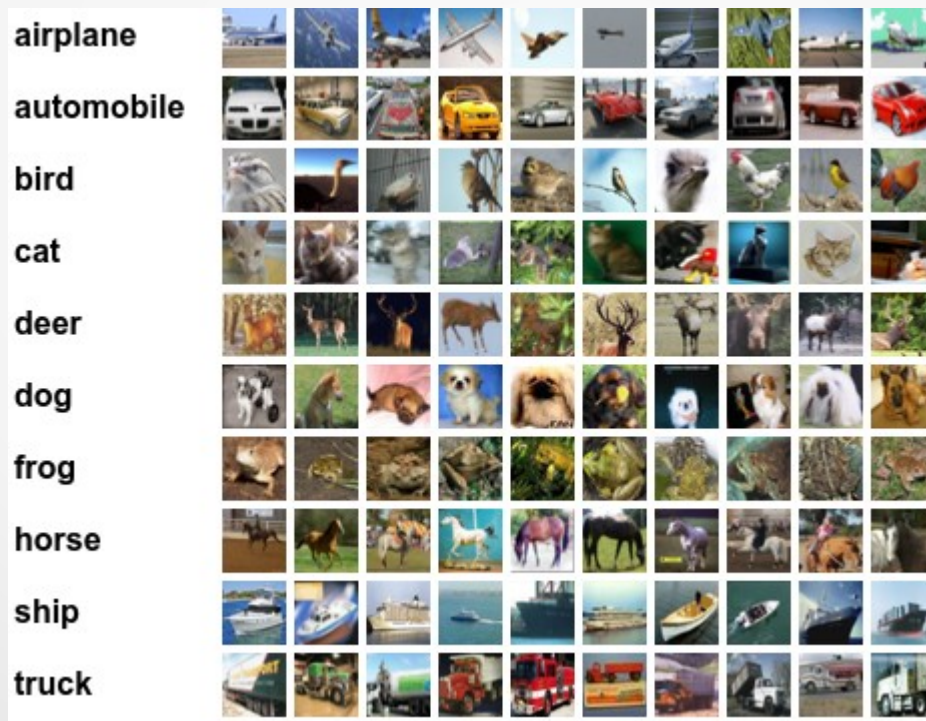




# Свёрточные нейросети

**CIFAR-10 (Canadian Institute for Advanced Research)**  
32x32 color, 60K images, 10 classes

<https://www.cs.toronto.edu/~kriz/cifar.html>



# Свёрточные нейросети

## ImageNet

14М изображений, 21К категорий

<http://www.image-net.org>



ILSVRC (ImageNet Large Scale Visual Recognition Challenge)





# Свёрточные нейросети

## **Pascal VOC (Visual Object Classes )**

11K images, 20 classes, 27K ROI annotated objects and 7K segmentations

<http://host.robots.ox.ac.uk/pascal/VOC/pubs/everingham10.pdf>



## Поиск объектов на изображении

git clone [https://github.com/mechanoid5/ml\\_lectorium.git](https://github.com/mechanoid5/ml_lectorium.git)

Борисов Е.С. О задаче поиска объекта на изображении.  
<http://mechanoid.su/cv-image-detector.html>

Борисов Е.С. Классификатор изображений на основе свёрточной сети. -- <http://mechanoid.su/ml-lenet.html>

Николенко С., Кадури́н А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей. - "Питер", 2018 г.

Fully Convolutional Networks for Semantic Segmentation  
<https://arxiv.org/pdf/1411.4038.pdf>

FCN — Fully Convolutional Network (Semantic Segmentation)  
<https://towardsdatascience.com/review-fcn-semantic-segmentation-eb8c9b50d2d1>