

# **Сверточные нейронные сети (Convolutional Neural Network - CNN)**

# Применение сверточных искусственных нейронных сетей в области компьютерного зрения

## 1. Классификация (Classification)

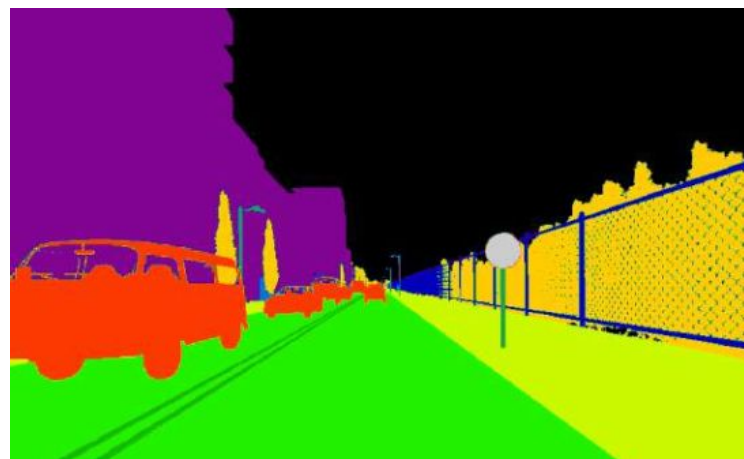
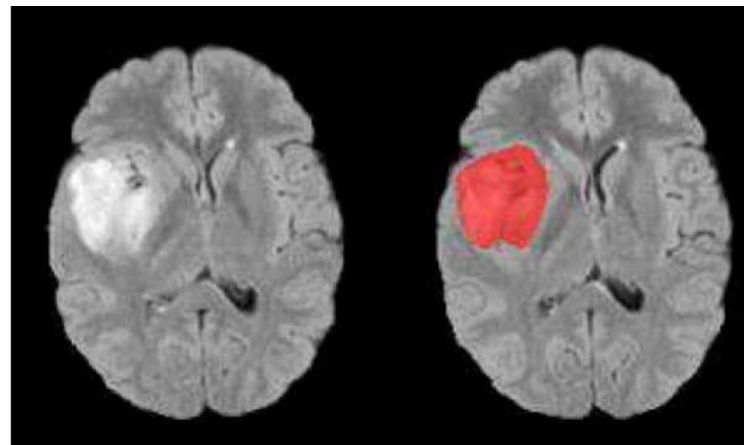


Cat - ?

## 2. Детектирование объектов (Object detection)



## 3. Семантическая сегментация (Semantic Segmentation)



# Операция свертки в CNN

Слой 1 (7x7)

0	1	1	1	0	0	0
0	0	1	1	1	0	0
0	0	0	1	1	1	0
0	0	0	1	1	0	0
0	0	1	1	0	0	0
0	1	1	0	0	0	0
1	1	0	0	0	0	0

**I**

1	0	1
0	1	0
1	0	1

Фильтр (3x3), stride = 1

**K**

1	4	3	4	1
1	2	4	3	3
1	2	3	4	1
1	3	3	1	1
3	3	1	1	0

Слой 2 (5x5)

**I \* K**

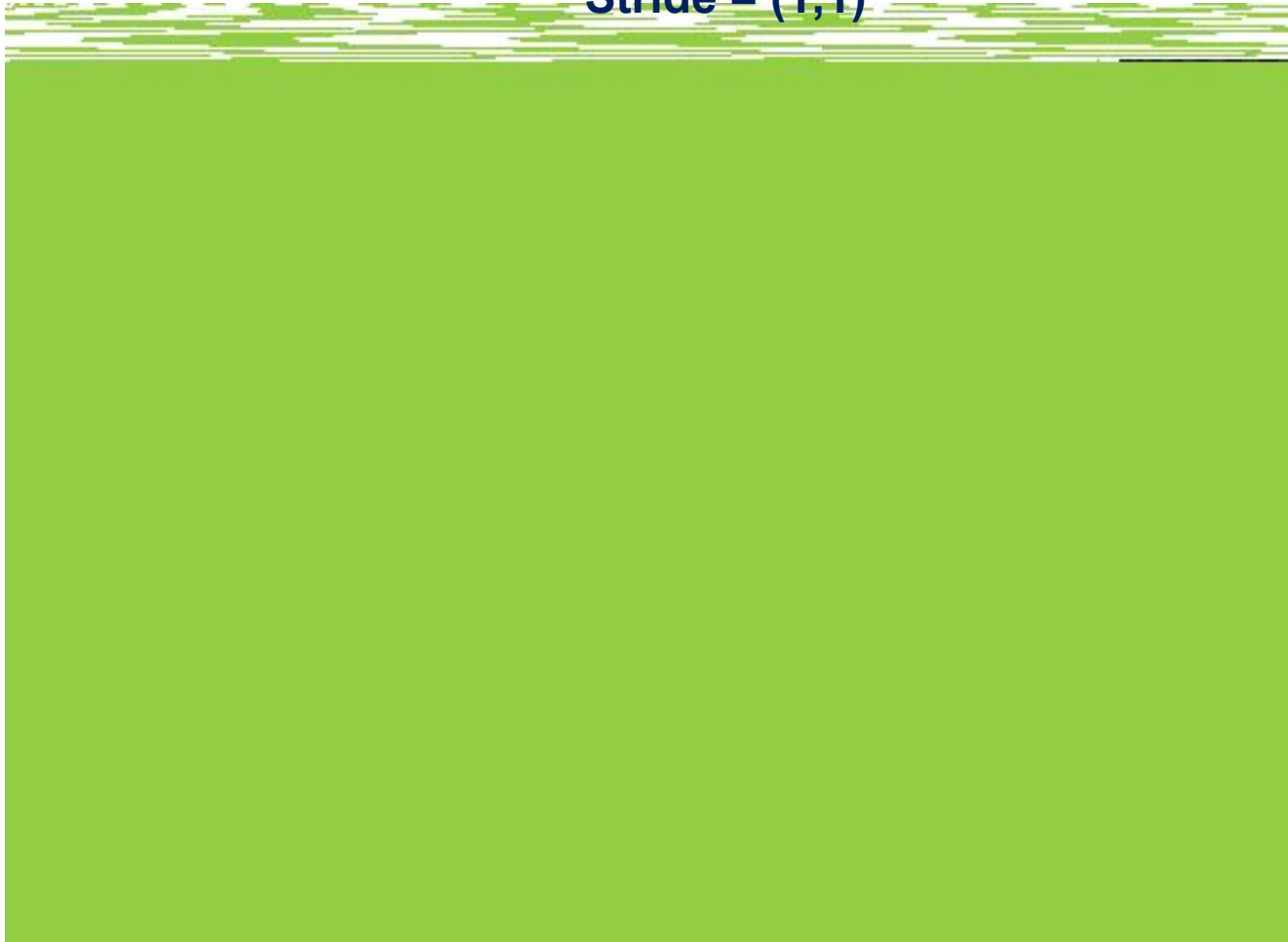
$$\text{conv}(I, K)_{x,y} = \sigma\left(b + \sum_{i=1}^h \sum_{j=1}^w \sum_{k=1}^d K_{ijk} \times I_{x+i-1,y+j-1,k}\right)$$

$b$  – смещение (bias),  $\sigma$  – функция активации

Stride – шаг движения фильтра (kernel)

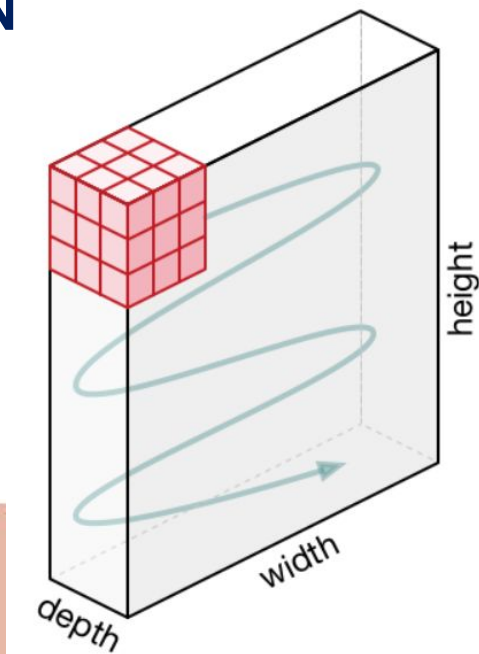
# Операция свертки в CNN

**Stride = (1,1)**



# Многоканальная свертка в CNN

Фильтр (N, N, depth)



Перемещение ядра

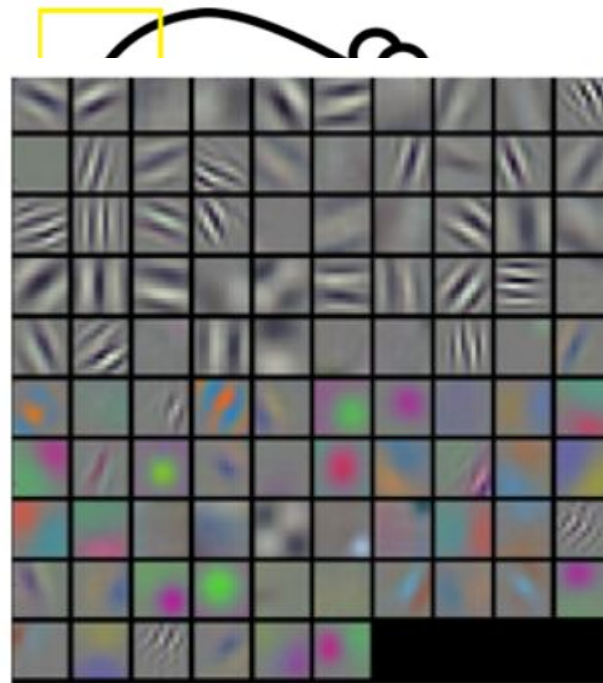


# Роль фильтров в CNN

Выделение характерных «паттернов» (прямых и изогнутых линий и т.п.) на изображении.

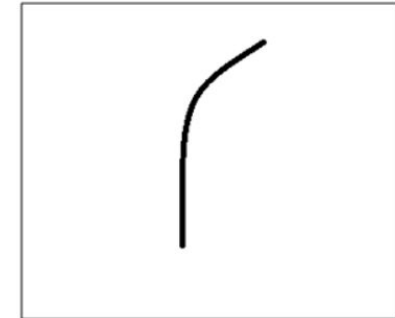


Original image



Visualizations of filters

фильтр



Visualization of a curve detector filter



Visualization of the receptive field

0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0

Pixel representation of the receptive field

0	0	30	0
0	30	0	0
30	0	0	0
30	0	0	0
30	0	0	0
30	0	0	0
30	0	0	0
0	0	0	0

Pixel representation of filter

Multiplication and Summation =  $(50 \cdot 30) + (50 \cdot 30) + (50 \cdot 30) + (20 \cdot 30) + (50 \cdot 30) = 6600$  (A large number!)

# Сверточные нейронные сети как прототип зрительной коры головного мозга

Хьюбел и Визель в 1962 году, в ходе эксперимента показали, что отдельные участки зрительной коры реагируют (или активируются) только при визуальном восприятии границ определенной ориентации. Например, некоторые нейроны активировались, когда воспринимали вертикальные границы, а некоторые — горизонтальные или диагональные. Хьюбел и Визель выяснили, что все эти нейроны сосредоточены в виде стержневой архитектуры и вместе формируют визуальное восприятие.

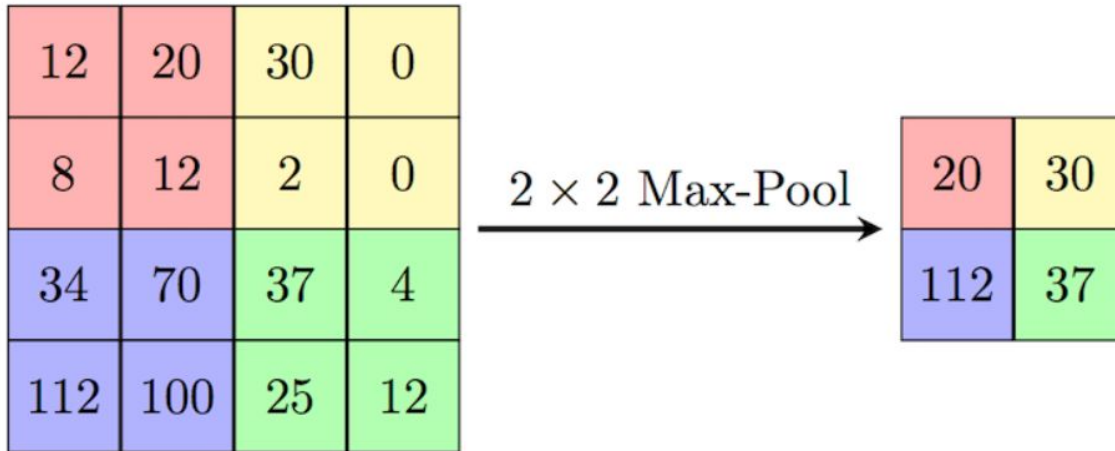
<https://youtu.be/Cw5PKV9Rj3o>

<https://youtu.be/IOHayh06LJ4>

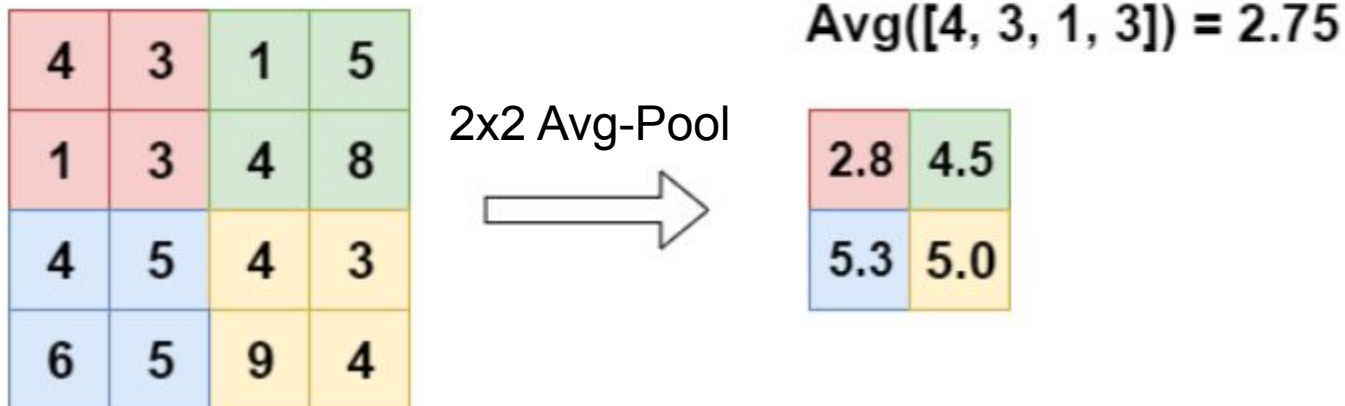


# Суб-дискретизация (Pooling)

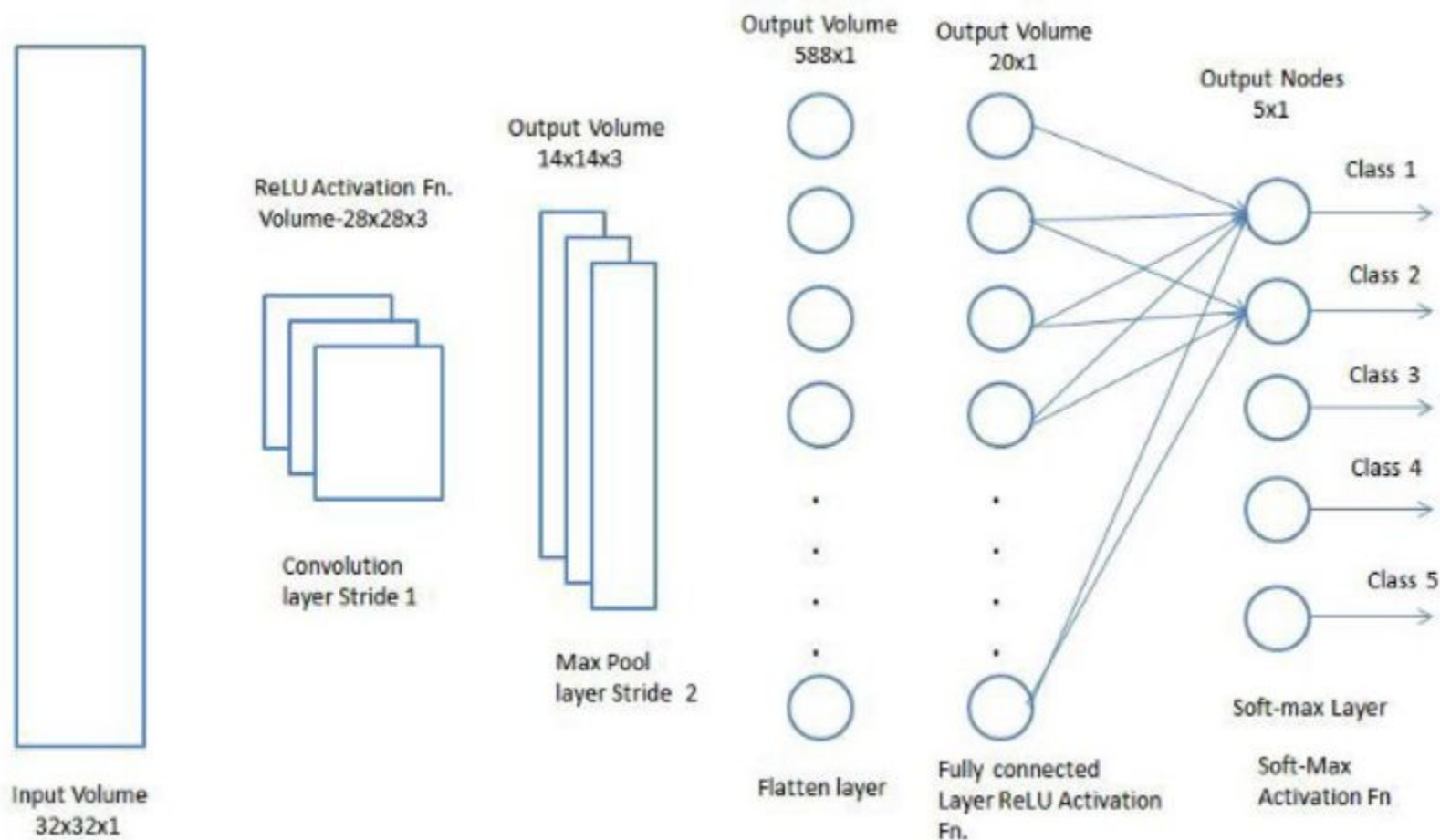
## 1) Max-Pooling



## 2) Average Pooling

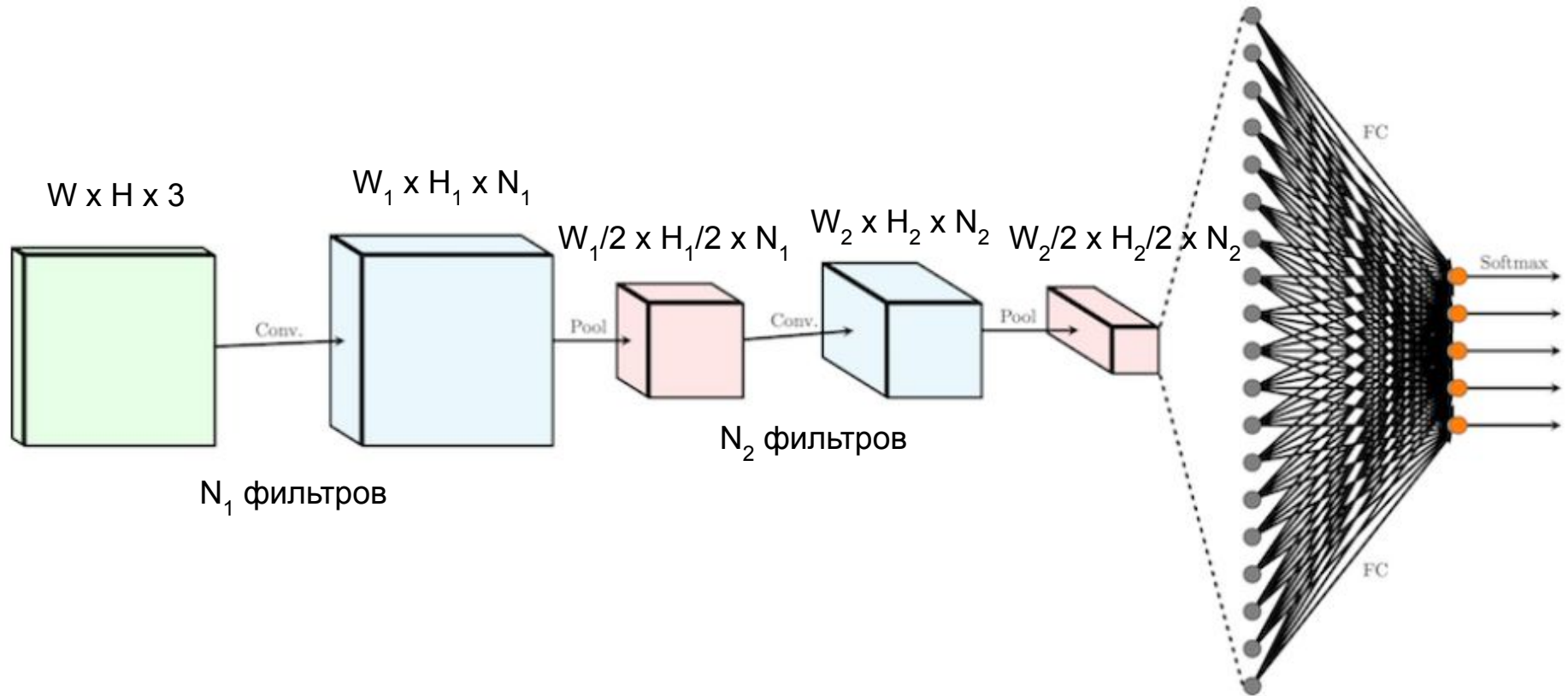


# Классификация в CNN при помощи полносвязных слоев нейронов



5 классов

# Типовая сверточная нейронная сеть для классификации объектов на изображении

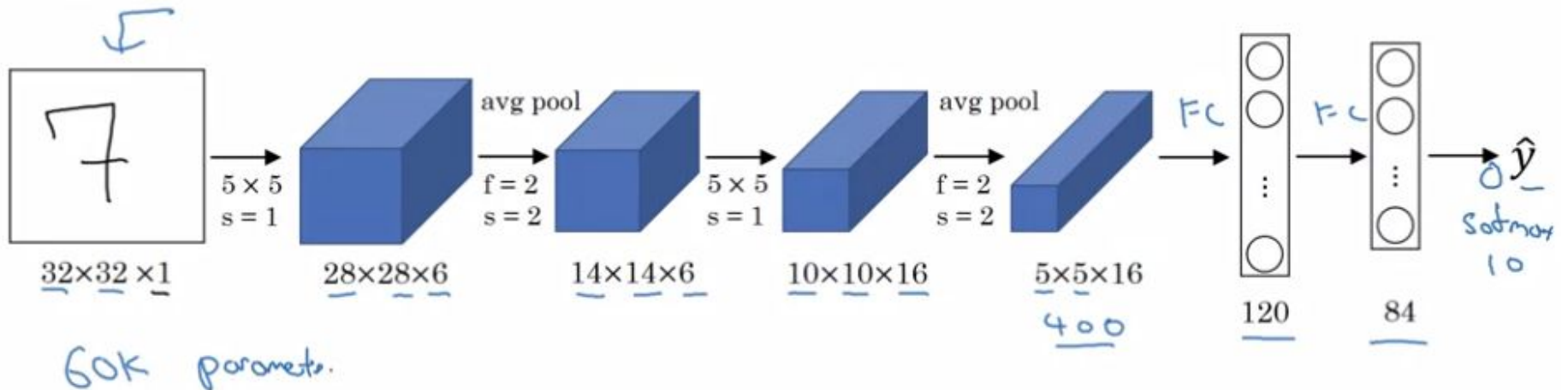


## **Преимущества сверточных нейронных сетей (CNN) над полносвязанными нейронными сетями (FCNN):**

- **Меньше обучаемых параметров между слоями**
- **Извлекая признаки из изображения, учитываются не только отдельные пиксели, но и окружающие пиксели (выявление неких паттернов на изображении)**

# Сверточная нейронная сеть LeNET-5 для задачи распознавания рукописных цифр

LeNet - 5



## Использованные материалы:

<https://habr.com/ru/post/309508/>

<https://habr.com/ru/post/510816/>

<https://habr.com/ru/company/wunderfund/blog/314872/>

<https://www.machinecurve.com/index.php/2020/01/30/what-are-max-pooling-average-pooling-global-max-pooling-and-global-average-pooling/>

<https://nuancesprog.ru/p/8307/>