

# Обучение нейронных сетей

Похоже ли обучение искусственной  
нейронной сети на обучение человека?

# Обучение «живых» нейронных сетей

- Нервная система формируется по программе заложенной в ДНК.
- «Дообучение» идет «без учителя» (обучение с подкреплением)
- Нет обратного распространения ошибок (?)
- Сеть не является полносвязной
- «Специализация» - выделяются подсети, решающие определенные задачи.

Вывод: для применения нейронной сети для решения задач, которые решает человек, нужно **усложнять архитектуру нейронной сети.**

## 2.3. Свёрточные нейронные сети (CNN)

- Происхождение – из компьютерного зрения (CV)
- Применение – для любой задачи, где между входными сигналами есть **пространственная упорядоченность**
  - Звук
  - Обработка естественного языка (NLP)
  - Анализ временных рядов
  - ...

# Как устроена зрительная система человека?

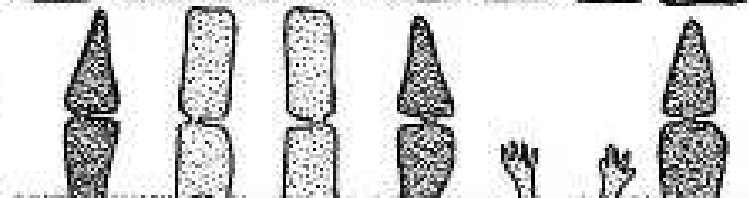
- Рецепторы сетчатки глаза - три слоя нервных клеток и два слоя синапсов:
  - около 5 млн нейронов-колбочек (цвет)
  - около 100 млн нейронов-палочек (ч/б)
- Зрительный нерв
  - от 770 000 до 1,7 миллиона аксонов, идущих от нейронов сетчатки глаза
- Зрительная кора
  - первичная зрительная кора (зона V1) – 140 миллионов нейронов
  - вторичная зрительная кора (зона V2) – имеет прямые и обратные связи с зоной V1
  - 30 других областей мозга связаны с зоной V1

Базальная мембрана



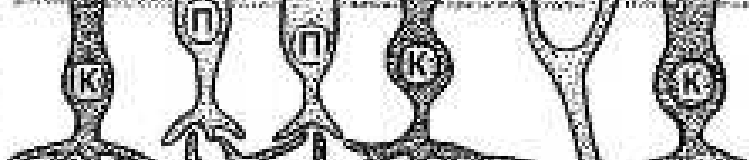
Пигментный эпителий

Наружные сегменты  
палочек и колбочек



Слой  
палочек и колбочек

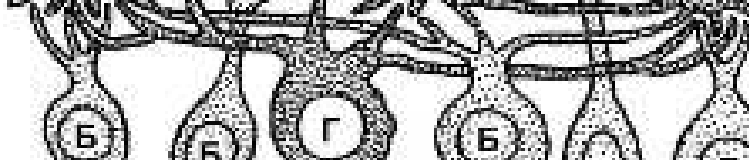
Внутренние сегменты  
палочек и колбочек



Наружная пограничная  
мембрана

Наружный  
ядерный слой

Наружный  
плексиформный слой



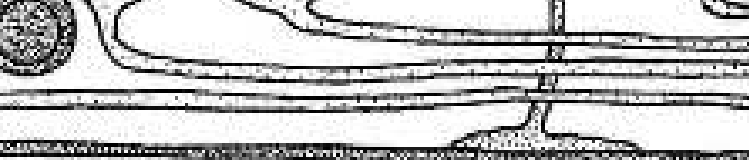
Внутренний  
ядерный слой

Внутренний  
плексиформный слой

Слой ганглиозных  
клеток

Слой нервных волокон

Внутренняя  
пограничная мембрана



Сосуд сетчатки



# Computer Vision (CV)

- фоторецепторы = пиксели
- палочки и колбочки = несколько входных каналов
- плексиморфные слои = свертки
- первичная зрительная кора = признаки
- вторичная зрительная кора = классификация

# Задачи CV

**Classification**



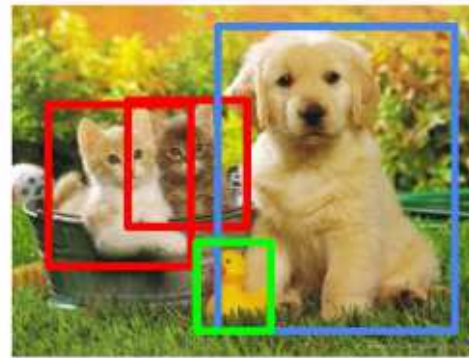
CAT

**Classification  
+ Localization**



CAT

**Object Detection**



CAT, DOG, DUCK

**Instance  
Segmentation**



CAT, DOG, DUCK

Single object

Multiple objects

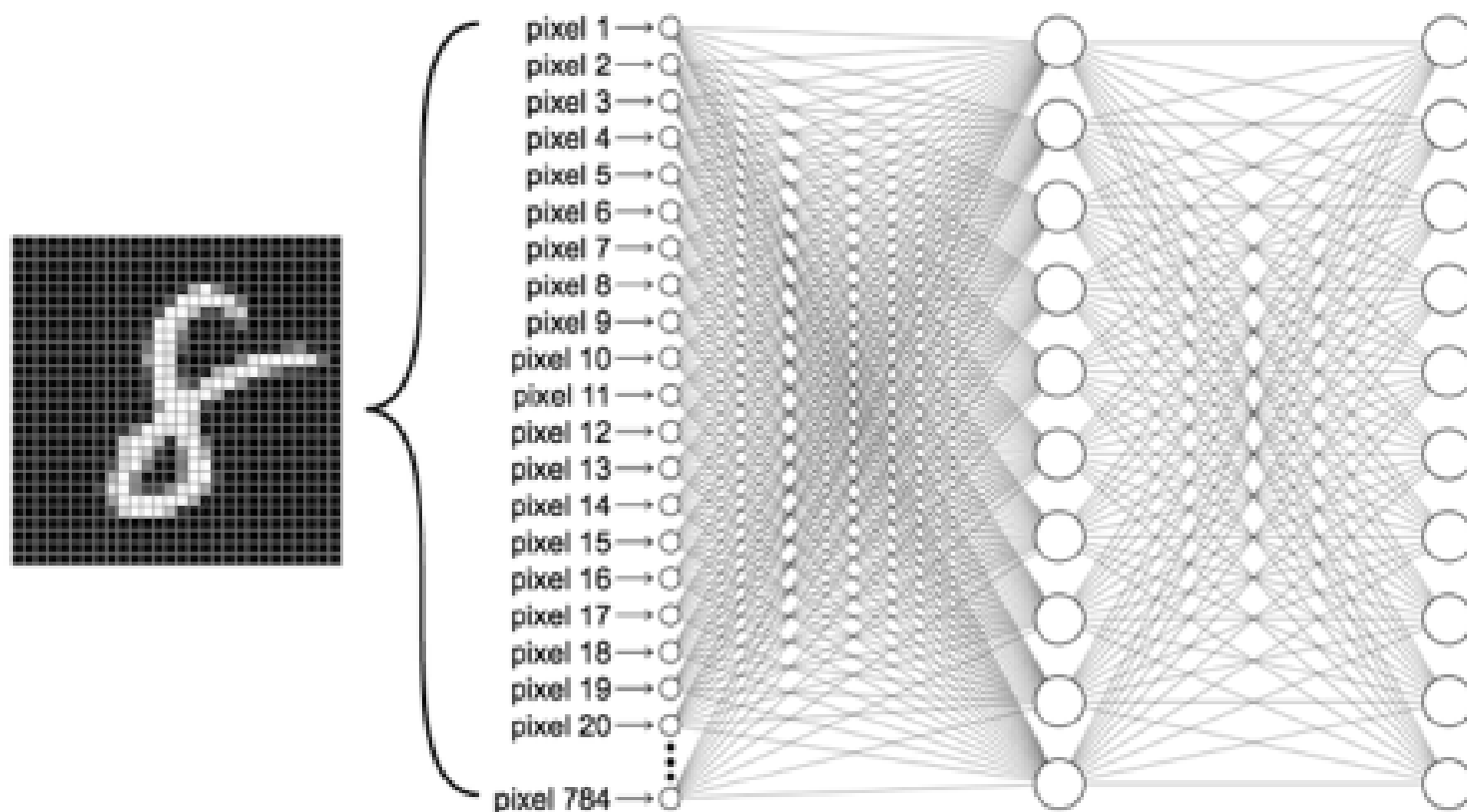
# MNIST

- Modified National Institute of Standards and Technology
- Содержит 70000 ч/б изображений рукописных цифр размером 28x28 пикселей
- Для калибровки и сопоставления методов распознавания изображений с помощью машинного обучения



[illegible]

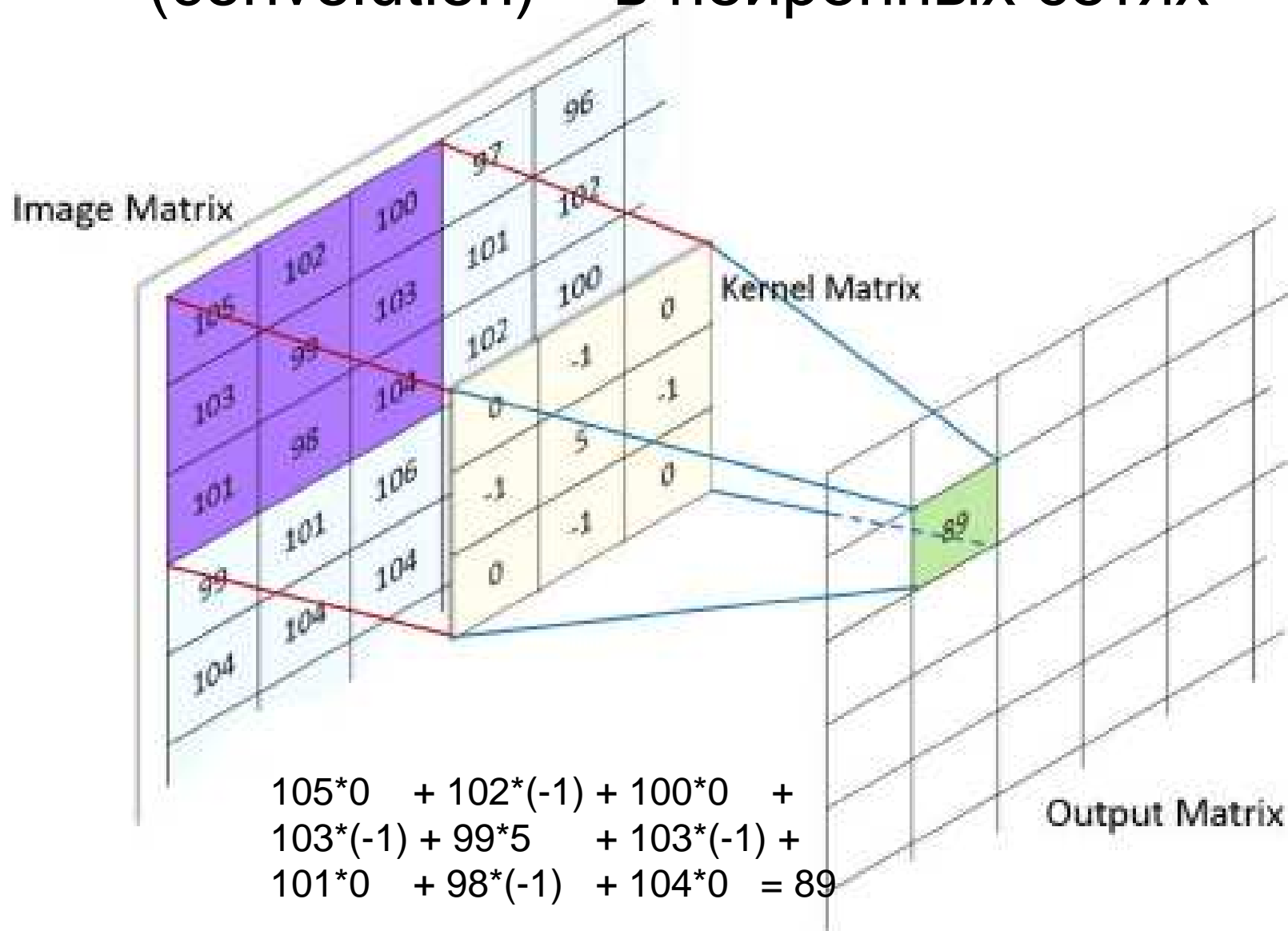
$28 \times 28 = 784$  входных сигнала?



# ImageNet

- База данных размеченных изображений
- С 2010 года – соревнования по классификации изображений
- Размер изображений:  $469 \times 387 \times 3 = 544\,509$  входных сигналов?
- Обычные полносвязные нейронные сети не применимы для решения таких задач

# Ян Лекун (1988): применение свёртки (convolution) – в нейронных сетях



# Операция свёртки применяется для обработки изображений

- Операция повторяется для каждого пикселя изображения
- Результат свертки нормируется, чтобы получилось снова изображение
- Изображение дополняется (экстраполируется) по краям (**padding**)
- На выходе получается изображение того же размера
- Это один из методов фильтрации – например гауссовский фильтр

# Примеры

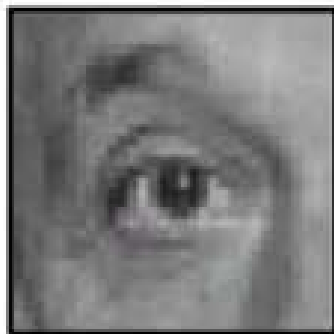


Original

0	0	0
0	1	0
0	0	0

?

# Примеры



Original

0	0	0
0	1	0
0	0	0

?



Original

0	0	0
0	1	0
0	0	0



Filtered  
(no change)

# Примеры



Original

0	0	0
0	0	1
0	0	0

?



# Примеры



Original

0	0	0
0	0	1
0	0	0

?



Original

0	0	0
0	0	1
0	0	0



Shifted left  
By 1 pixel

# Примеры



Original

 $\frac{1}{9}$ 

1	1	1
1	1	1
1	1	1

?

# Примеры



Original

 $\frac{1}{9}$ 

1	1	1
1	1	1
1	1	1

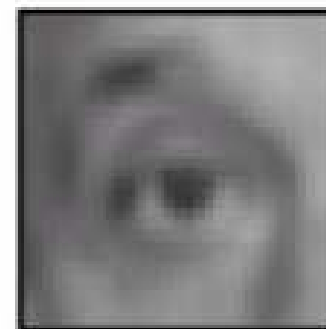
?



Исходное

 $\frac{1}{9}$ 

1	1	1
1	1	1
1	1	1

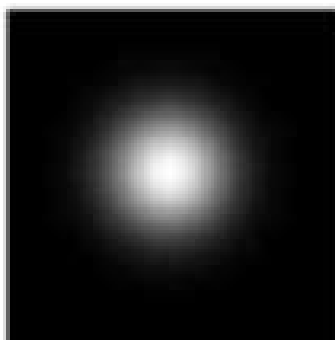
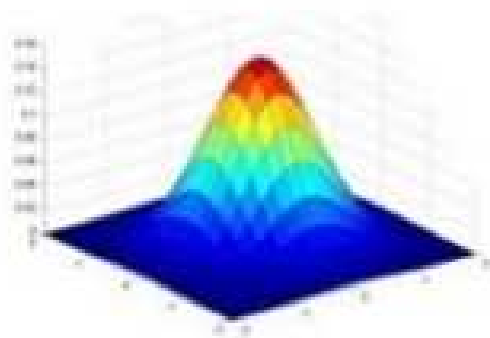


Результат

# Фильтр Гаусса

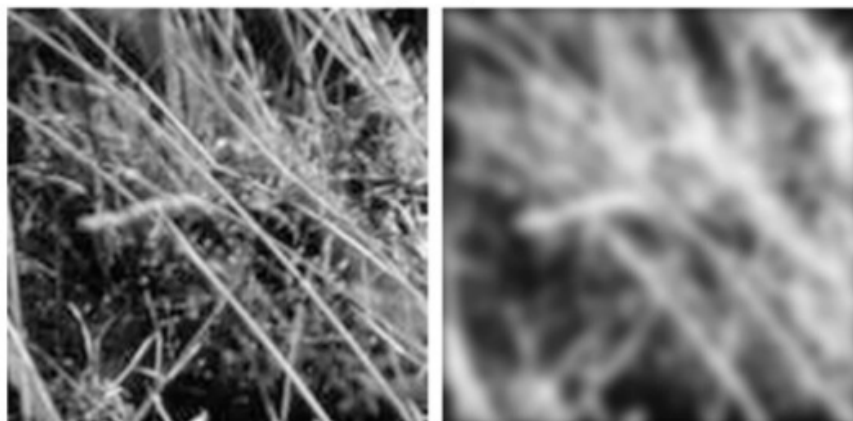
Задаём вес пикселей при  
сглаживании с учётом  
близости к центру:

$$G_{\sigma} = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$



0.003	0.013	0.022	0.013	0.003
0.013	0.059	0.097	0.059	0.013
0.022	0.097	0.159	0.097	0.022
0.013	0.059	0.097	0.059	0.013
0.003	0.013	0.022	0.013	0.003

5 x 5,  $\sigma = 1$



# Фильтр для повышения резкости

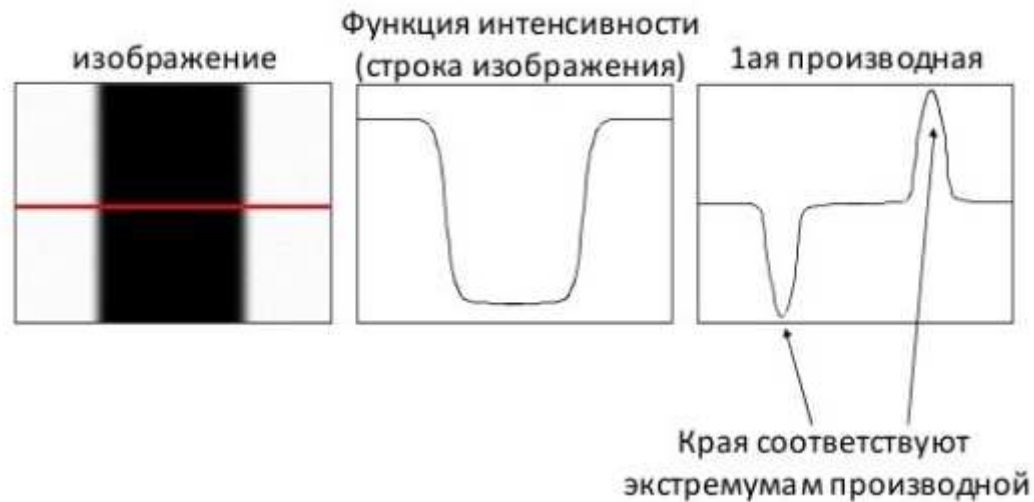
Ядро  
свертки

$$\frac{1}{10} \cdot \begin{bmatrix} -1 & -2 & -1 \\ -2 & 22 & -2 \\ -1 & -2 & -1 \end{bmatrix}$$



# Выделение границ

Граница – это резкое изменение функции интенсивности.  
 Может быть найдено с помощью разности функции интенсивности



$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Робертса

Превитт

Собеля

# Свёртка = слой нейронной сети

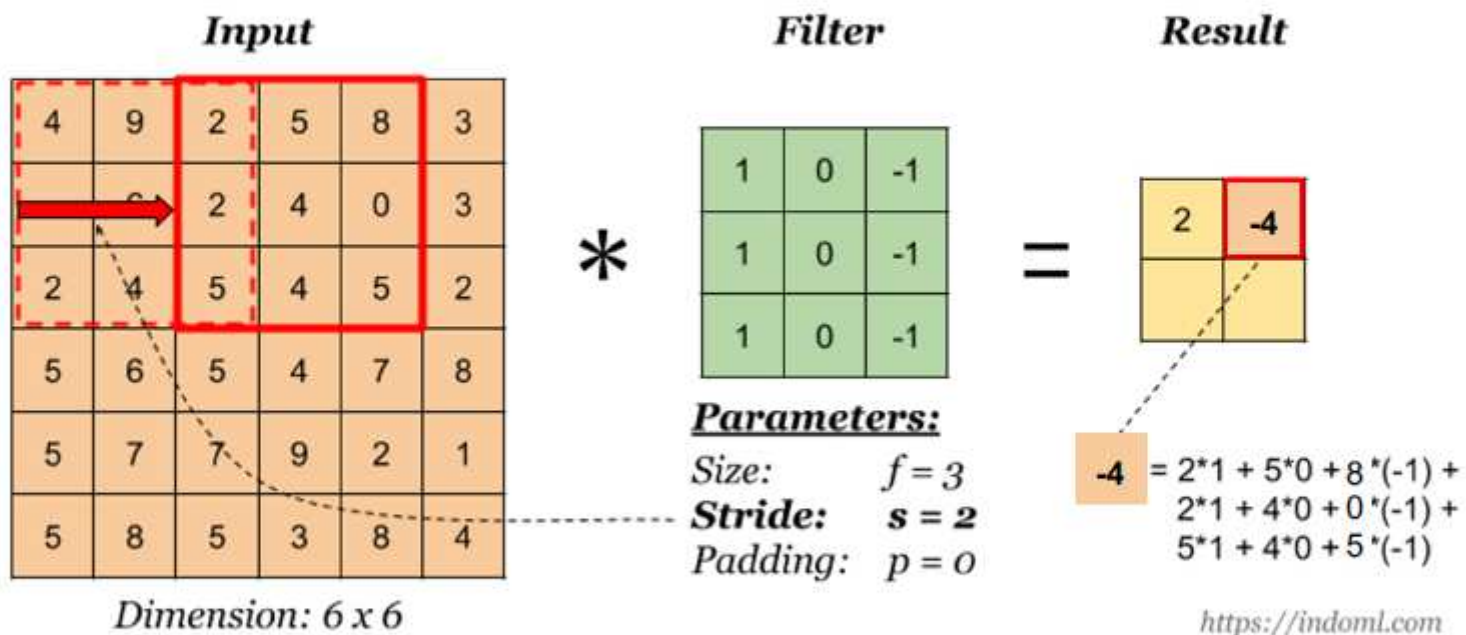
- Элемент свёртки – это «вырожденный» нейрон, у которого есть связи только с несколькими входами
- Входы расположены на небольшом расстоянии друг от друга
- Нейроны свёртки повторяют геометрию входных сигналов
- Свертки идут слоями – в результате нейрон свертки более глубокого слоя охватывает более широкое поле исходного изображения
- Веса нейрона обучаются
- Веса нейронов одного сверточного слоя одинаковые

# Шаг свёртки (stride)

Stride = 2

$$H_{out} = (H - F) // k + 1 = 2$$

$$W_{out} = (W - F) // k + 1 = 2$$



H=6, W=6, C=1  
(H, W, C)

F=3  
(F, F, C)



## Обрамление нулями (padding)

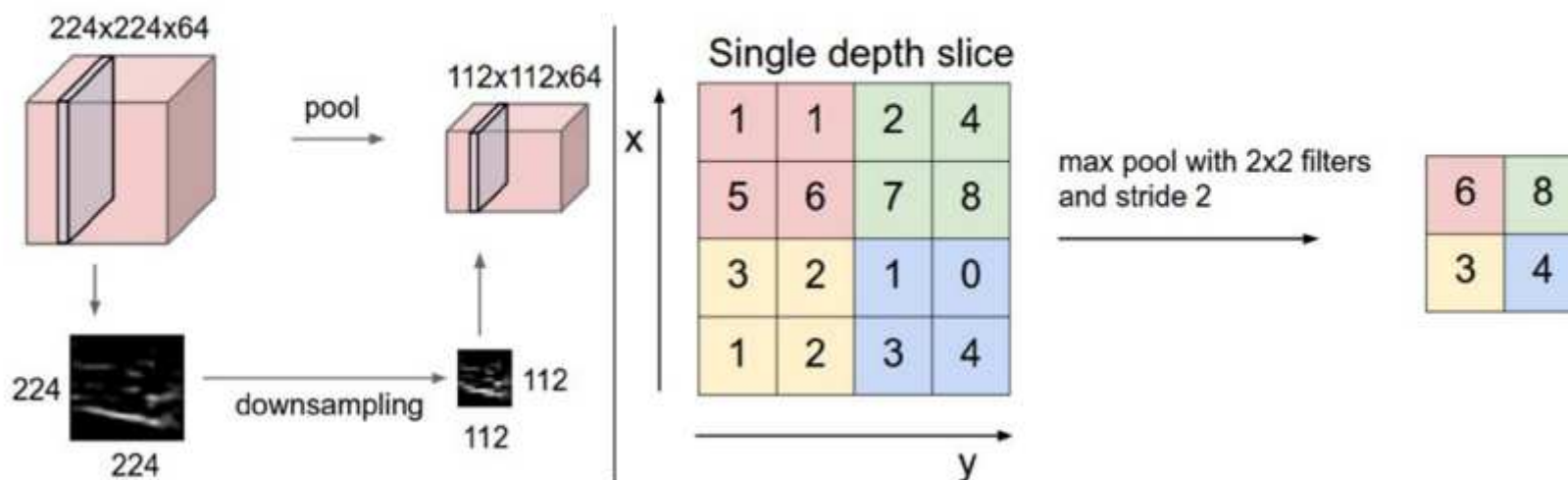
0 <sub>2</sub>	0 <sub>0</sub>	0 <sub>1</sub>	0	0	0	0
0 <sub>1</sub>	2 <sub>0</sub>	2 <sub>0</sub>	3	3	3	0
0 <sub>0</sub>	0 <sub>1</sub>	1 <sub>1</sub>	3	0	3	0
0	2	3	0	1	3	0
0	3	3	2	1	2	0
0	3	3	0	2	3	0
0	0	0	0	0	0	0

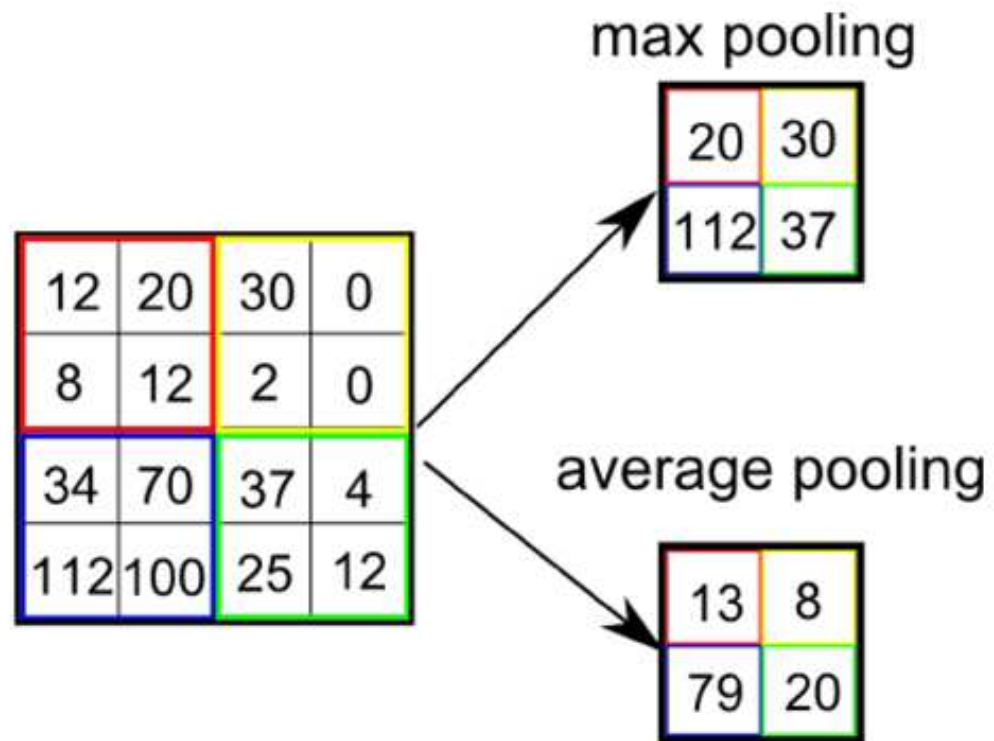
1	6	5
7	10	9
7	10	8

# Размер результата свёртки

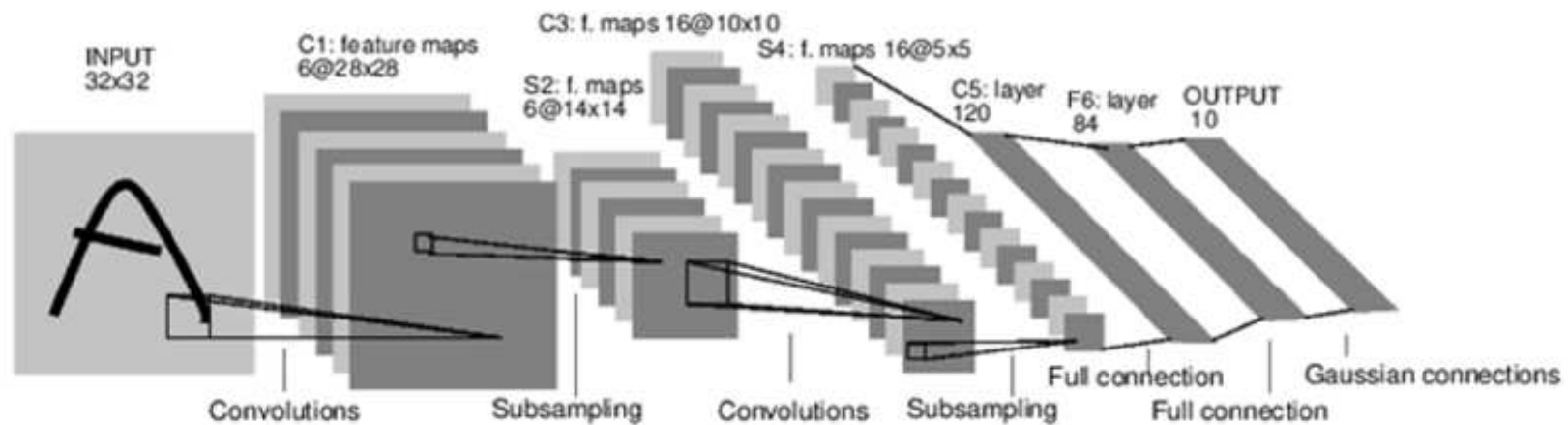
- Размер картинки:  $(H, W, 3)$
- Размер фильтра:  $(F, F, 3)$
- $\text{Stride} = S$
- $\text{Padding} = P$
- Размер тензора-результата
- $(H_{\text{out}}, W_{\text{out}}, C_{\text{out}})$ :
  - $H_{\text{out}} = (H - F + 2P) // S + 1$
  - $W_{\text{out}} = (W - F + 2P) // S + 1$
  - $C_{\text{out}} = 1$

# Pooling слой





# Свёрточная нейронная сеть



A Full Convolutional Neural Network (LeNet)

- Комбинация сверточных слоев, pooling-слоев, и полносвязных слоев
- Метод обратного распространения ошибок также применим для обучения весов свёрточного слоя

# Пример

