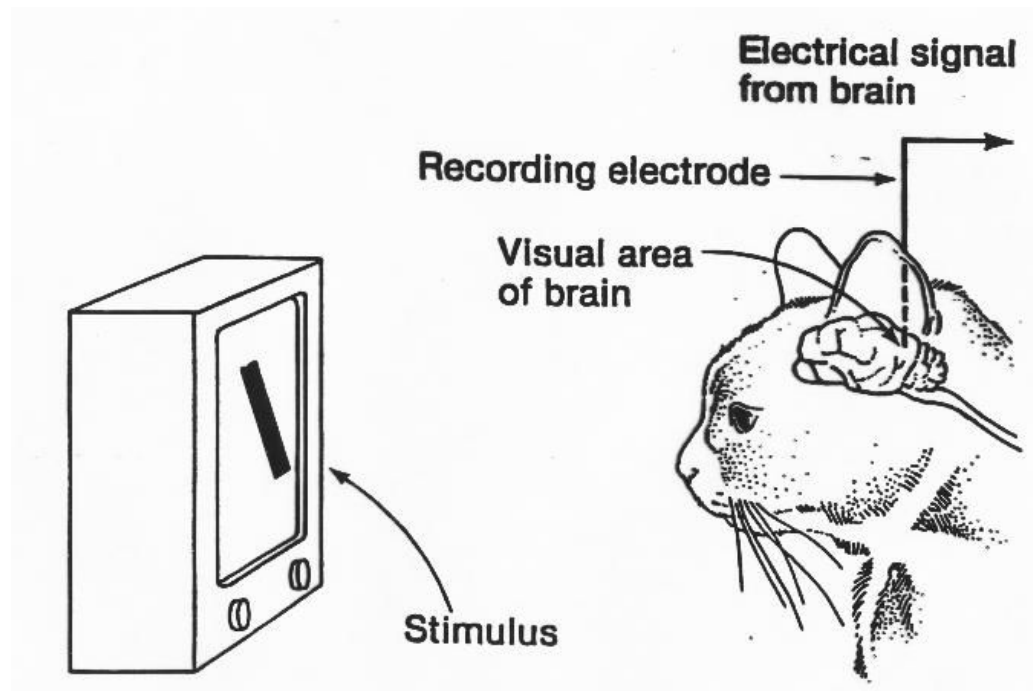


Сверточные нейронные сети

Студент 172 группы, Федоров Павел

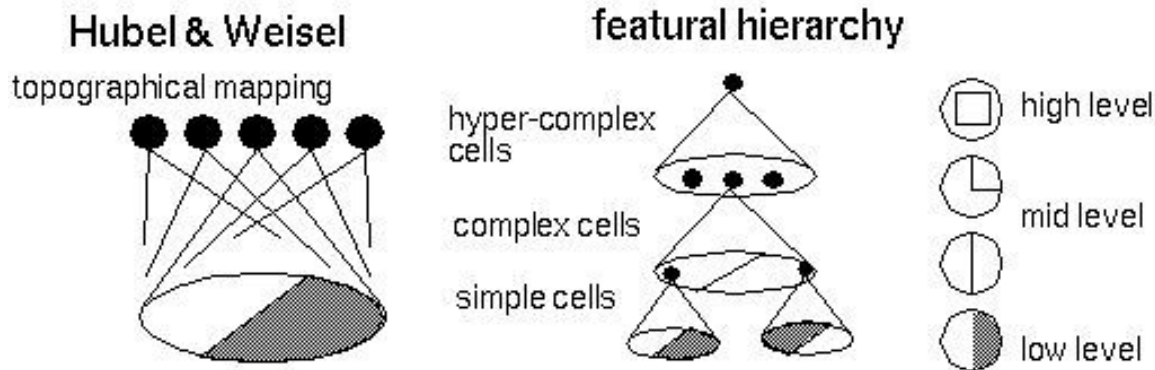
Идея сверточных сетей



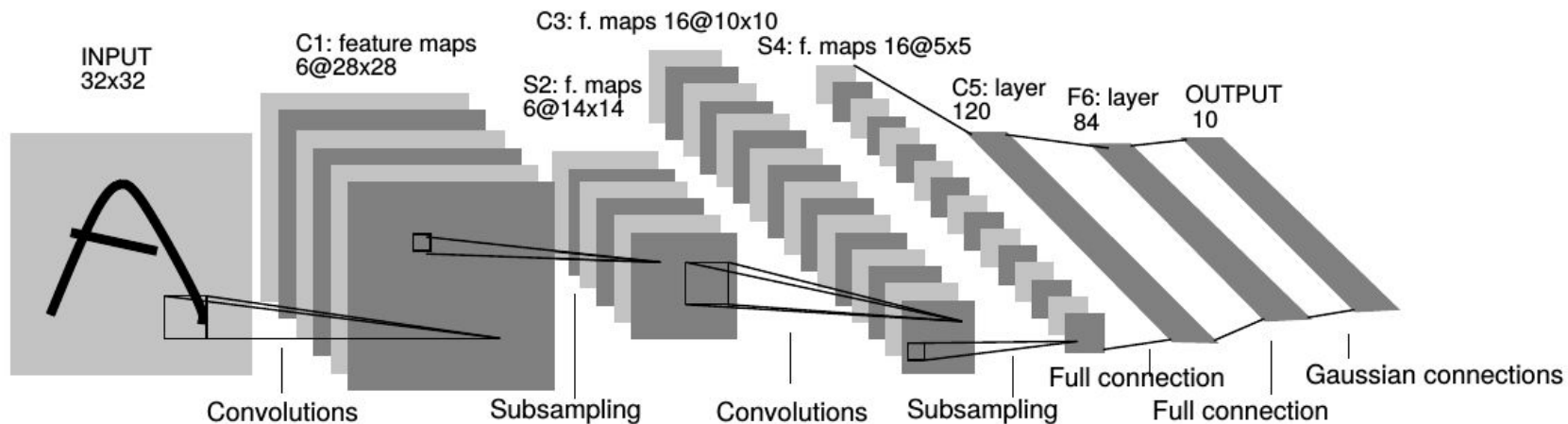
Эксперимент Дэвида Хьюбела и Торстена Визеля (1962)

Результаты эксперимента

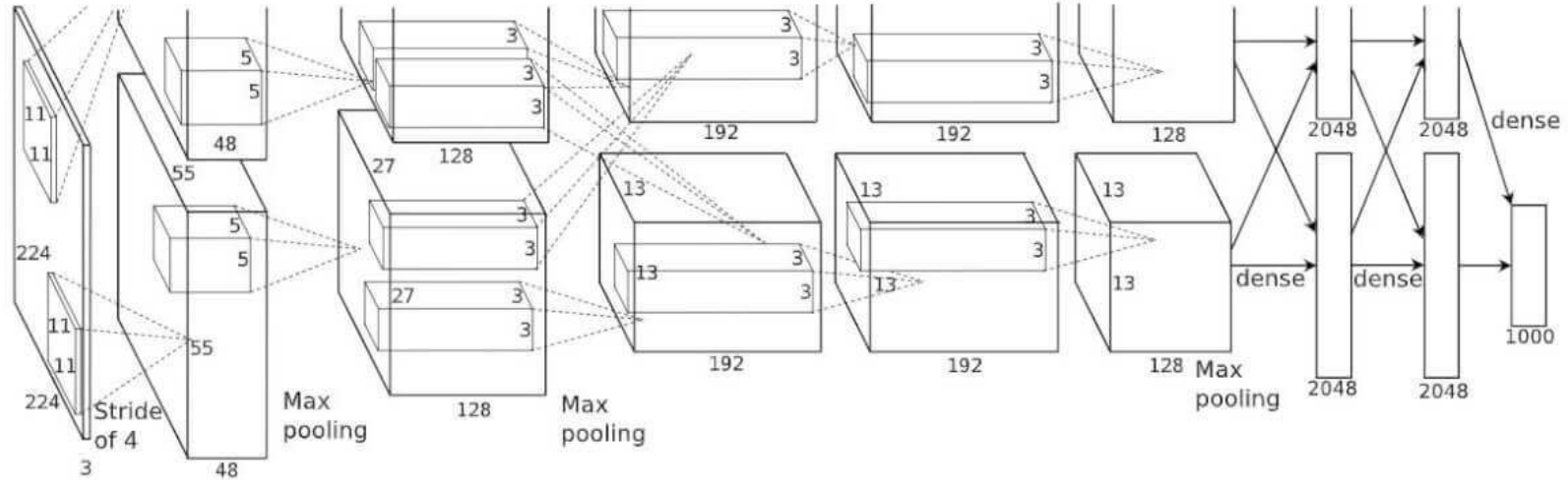
- Соседние нейроны обрабатывают сигналы с соседних областей сетчатки;
- Нейроны образуют иерархическую структуру;
- Нейроны организованы в так называемые колонки — вычислительные блоки, которые трансформируют и передают информацию от уровня к уровню.



Ян ЛеКун и LeNet-5 (1998)

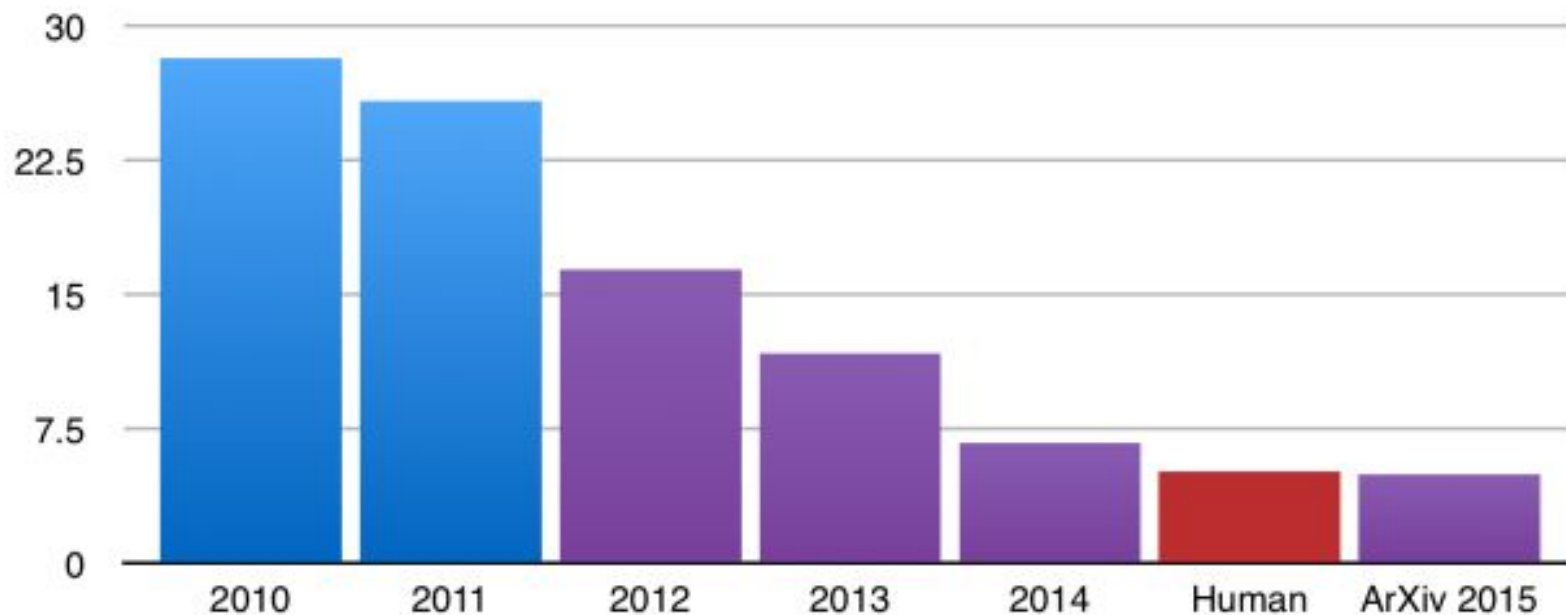


LeNet + теория обучения + большие данные + *железо*
= AlexNet (2012)



Успех сверточных нейронных сетей

ILSVRC top-5 error on ImageNet



Задача классификации

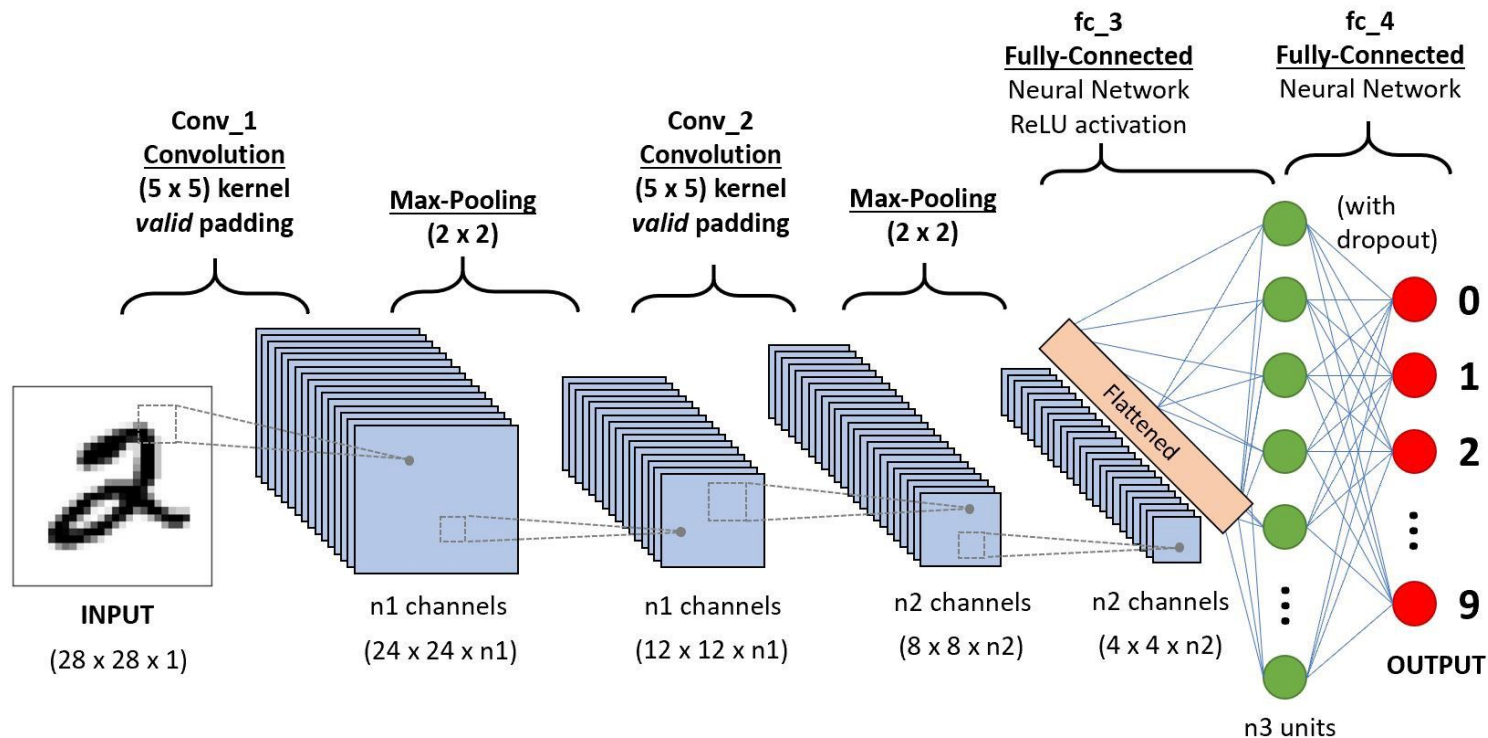


What We See

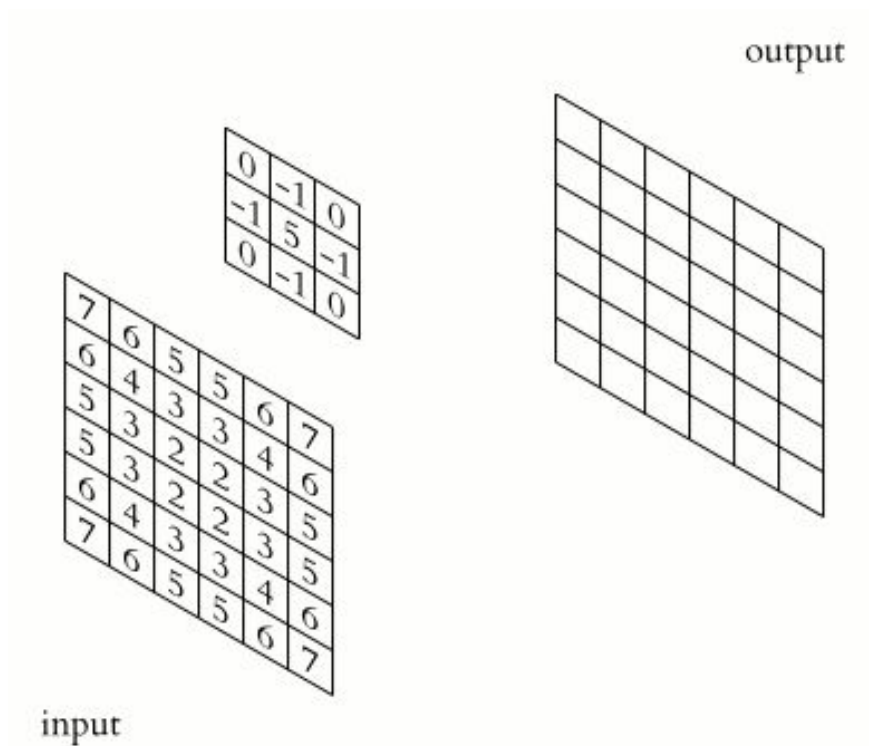
```
08 02 22 97 38 15 00 40 00 75 04 05 07 78 52 12 50 77 91 08
49 49 99 40 17 81 18 57 60 87 17 40 98 43 69 48 04 56 62 00
81 49 31 73 55 79 14 29 93 71 40 67 53 88 30 03 49 13 36 65
52 70 95 23 04 60 11 42 69 24 68 56 01 32 56 71 37 02 36 91
22 31 16 71 51 67 63 89 41 92 36 54 22 40 40 28 66 33 13 80
24 47 32 60 99 03 45 02 44 75 33 53 78 36 84 20 35 17 12 50
32 98 81 28 64 23 67 10 26 38 40 67 59 54 70 66 18 38 64 70
67 26 20 68 02 62 12 20 95 63 94 39 63 08 40 91 66 49 94 21
24 55 58 05 66 73 99 26 97 17 78 78 96 83 14 88 34 89 63 72
21 36 23 09 75 00 76 44 20 45 35 14 00 61 33 97 34 31 33 95
78 17 53 28 22 75 31 67 15 94 03 80 04 62 16 14 09 53 56 92
16 39 05 42 96 35 31 47 55 58 88 24 00 17 54 24 36 29 85 57
86 56 00 48 35 71 89 07 05 44 44 37 44 60 21 58 51 54 17 58
19 80 81 68 05 94 47 69 28 73 92 13 86 52 17 77 04 89 55 40
04 52 08 83 97 35 99 16 07 97 57 32 16 26 26 79 33 27 98 66
88 36 68 87 57 62 20 72 03 46 33 67 46 55 12 32 63 93 53 69
04 42 16 73 38 25 39 11 24 94 72 18 08 46 29 32 40 62 76 36
20 69 36 41 72 30 23 88 34 62 99 69 82 67 59 85 74 04 36 16
20 73 35 29 78 31 90 01 74 31 49 71 48 86 81 16 23 57 05 54
01 70 54 71 83 51 54 69 16 92 33 48 61 43 52 01 89 19 67 48
```

What Computers See

Структура



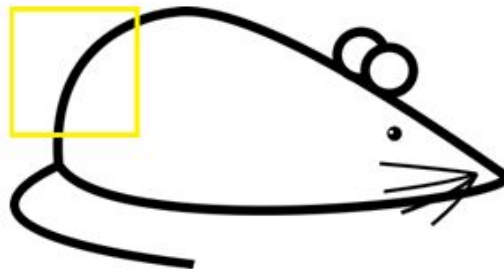
Операция свертки



Пример ядра свертки



Original image



Visualization of the filter on the image



Visualization of the receptive field

0	0	0	0	0	0	30
0	0	0	0	50	50	50
0	0	0	20	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0

Pixel representation of the receptive field

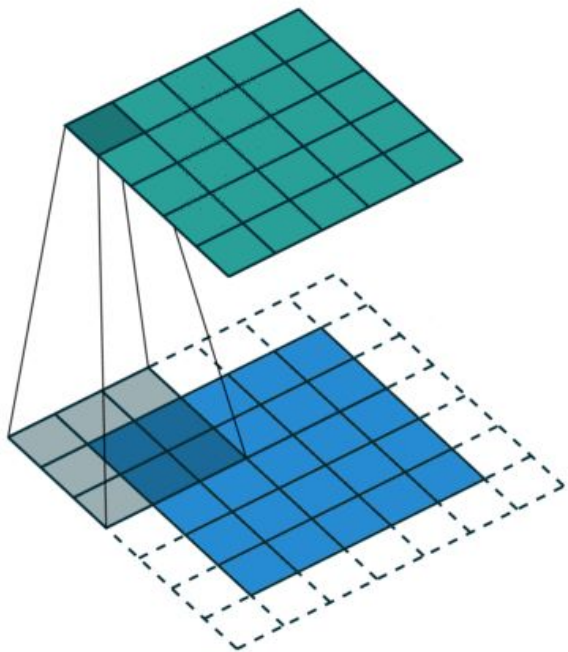
*

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

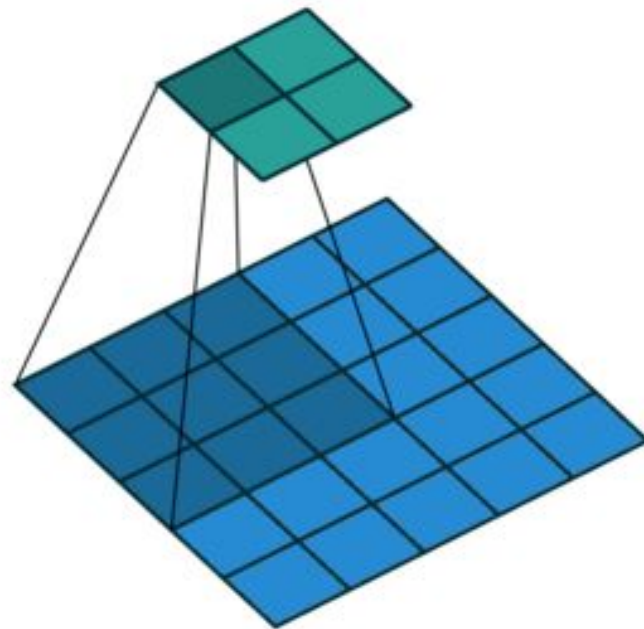
Pixel representation of filter

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

Padding и Striding



Padding (добавление рамки)



Striding (размер шага ядра)

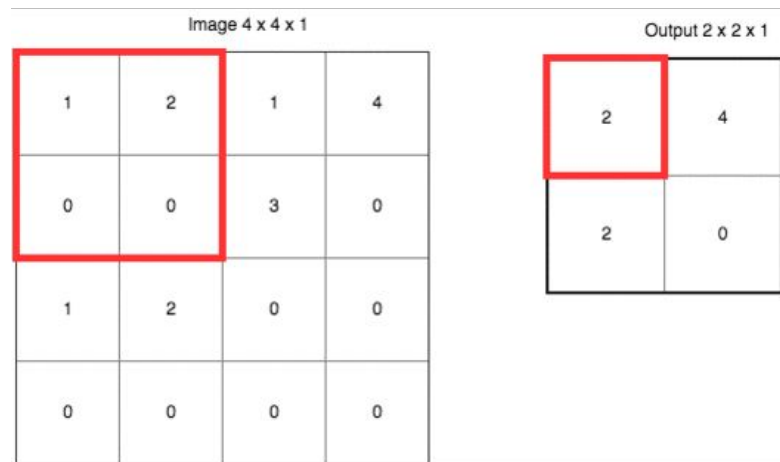
Считаем размер выхода сверточного слоя

- Вход: тензор $H_{in} \cdot W_{in} \cdot F_{in}$
- Выход: тензор $H_{out} \cdot W_{out} \cdot F_{out}$
- 4 гиперпараметра:
 - F: число фильтров
 - H, W : пространственный размер фильтров
 - S: шаг
 - P: количество заполнения нулями
- Соотношение размеров входа и выхода:

$$\begin{aligned}H_{out} &= (H_{in} - H_k + 2P)/S + 1 \\W_{out} &= (W_{in} - W_k + 2P)/S + 1 \\F_{out} &= F\end{aligned}$$

Даунсемплинг

- Замена выхода нейрона статистикой, подсчитанной по его соседям
 - max pooling
 - (weighted) average pooling
- Обеспечивает приблизительную инвариантность выхода к малому переносу входа
- Наличие признака важнее, чем его точная позиция



Пример

Max pooling

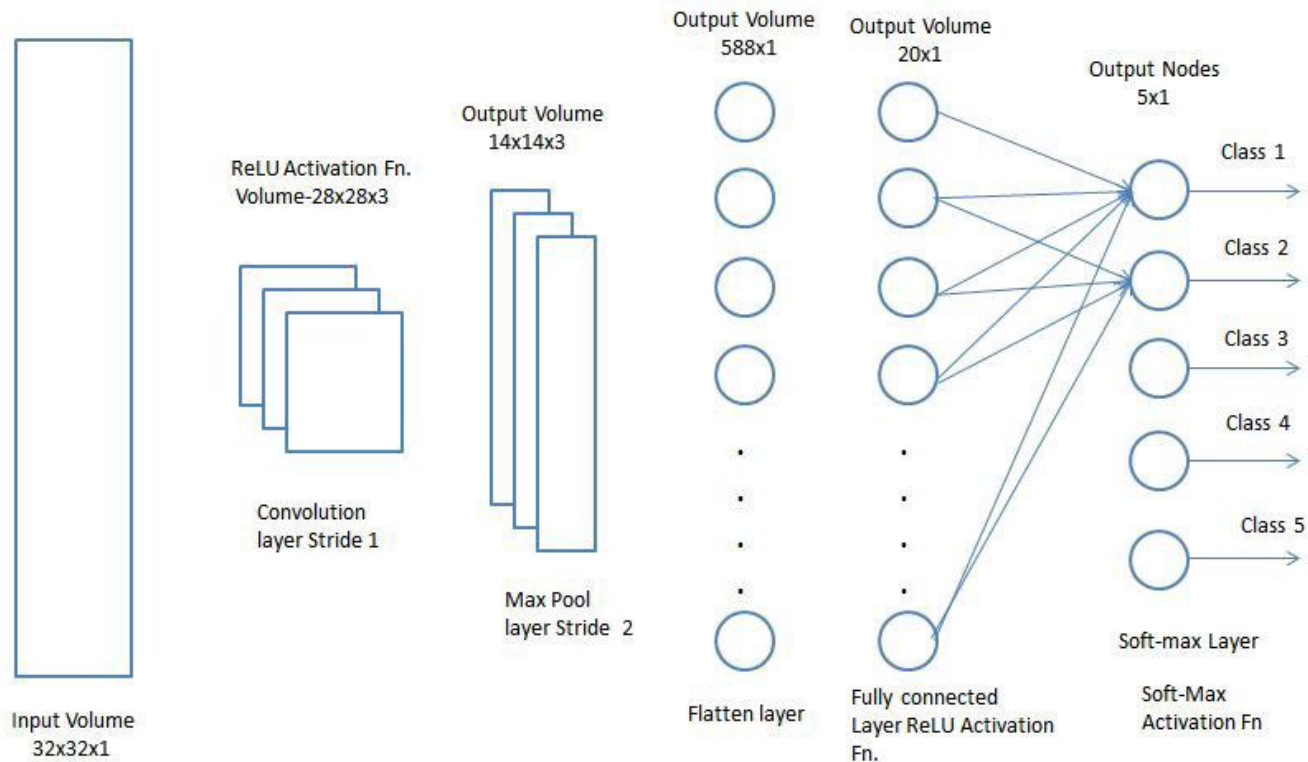


Считаем размер выхода pooling слоя

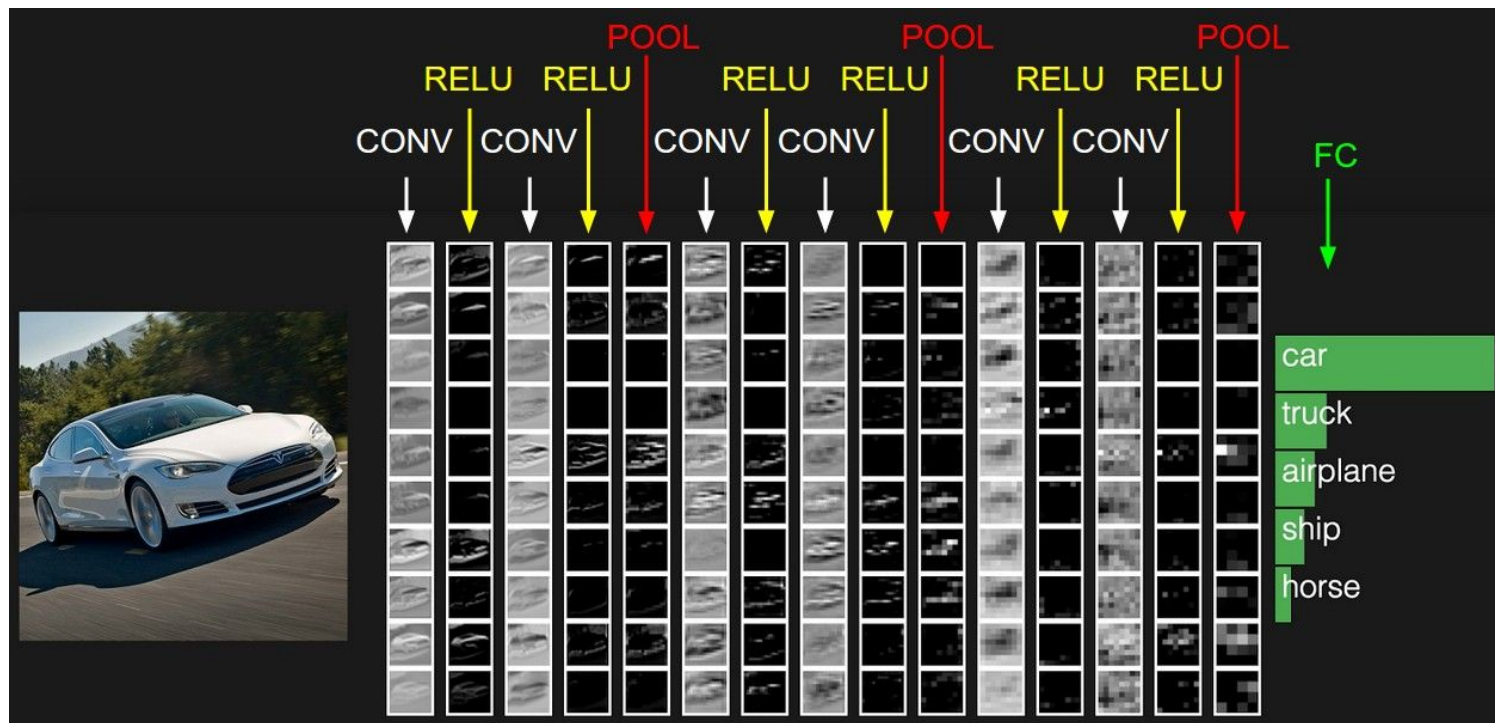
- Вход: тензор $H_{in} \cdot W_{in} \cdot F_{in}$
- Выход: тензор $H_{out} \cdot W_{out} \cdot F_{out}$
- 3 гиперпараметра:
 - F: пространственный размер объединения
 - S: шаг объединения
- Соотношение размеров входа и выхода:

$$\begin{aligned}H_{out} &= (H_{in} - F)/S + 1 \\W_{out} &= (W_{in} - F)/S + 1 \\F_{out} &= F\end{aligned}$$

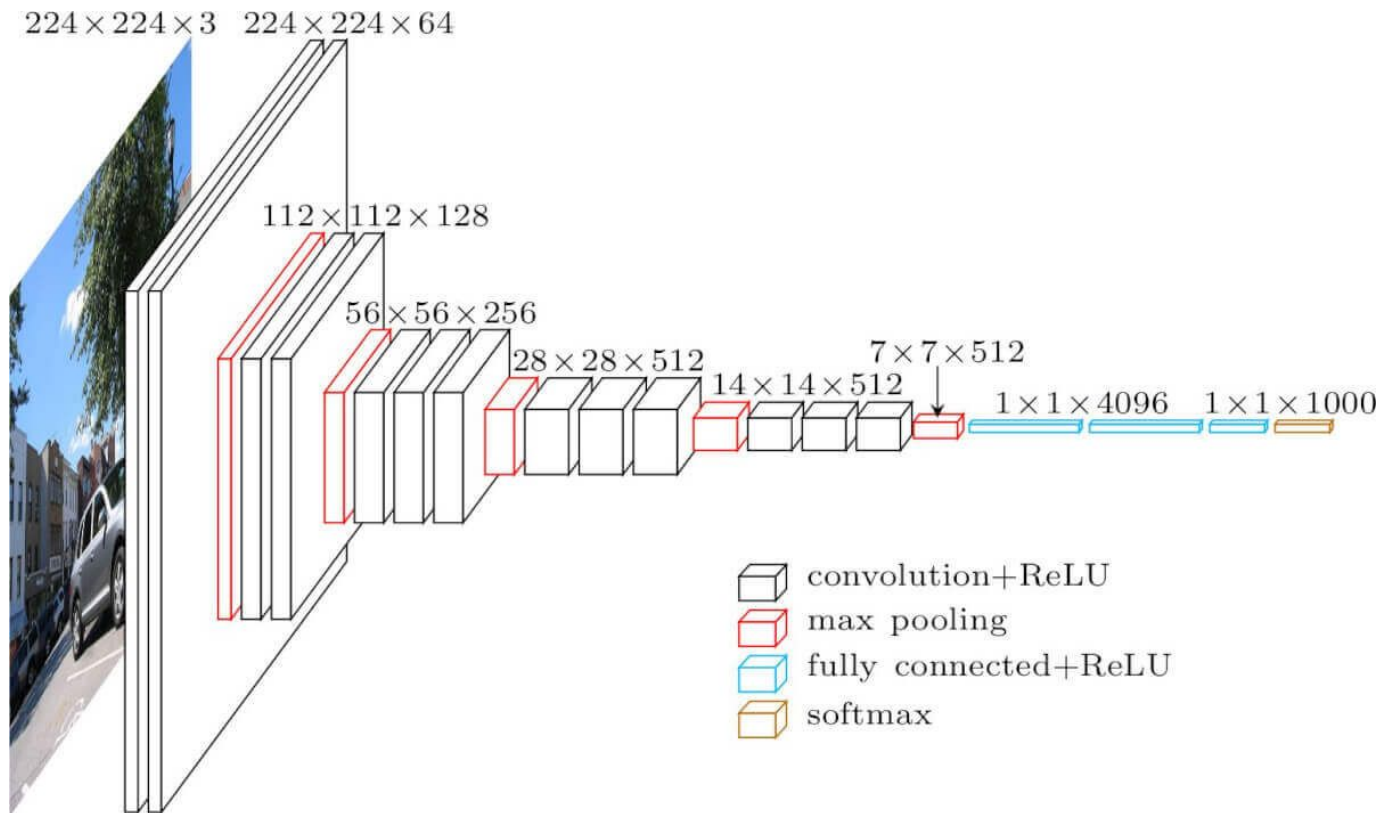
Полносвязный слой



Пример работы сверточной сети



VGG-16



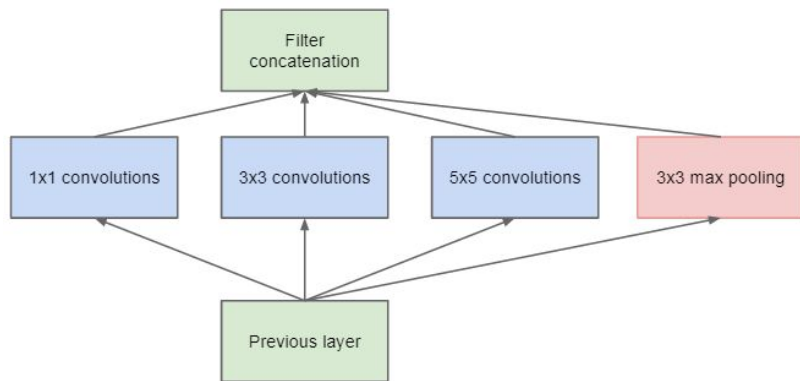
Недостатки VGG

- Очень медленная скорость обучения.
- Сама архитектура сети весит слишком много (появляются проблемы с диском и пропускной способностью)

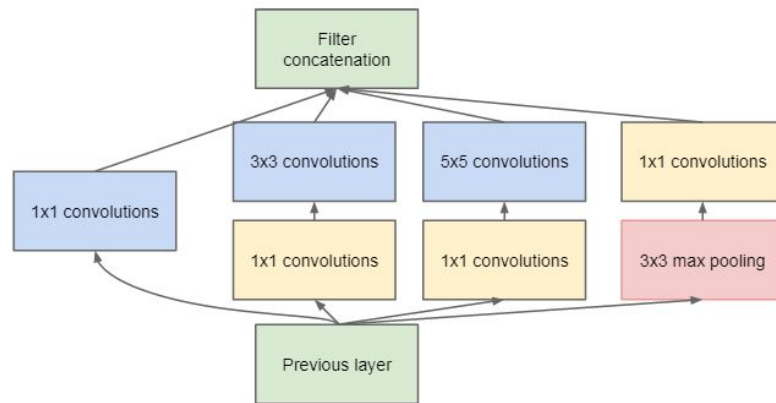
Inception



Inception слой

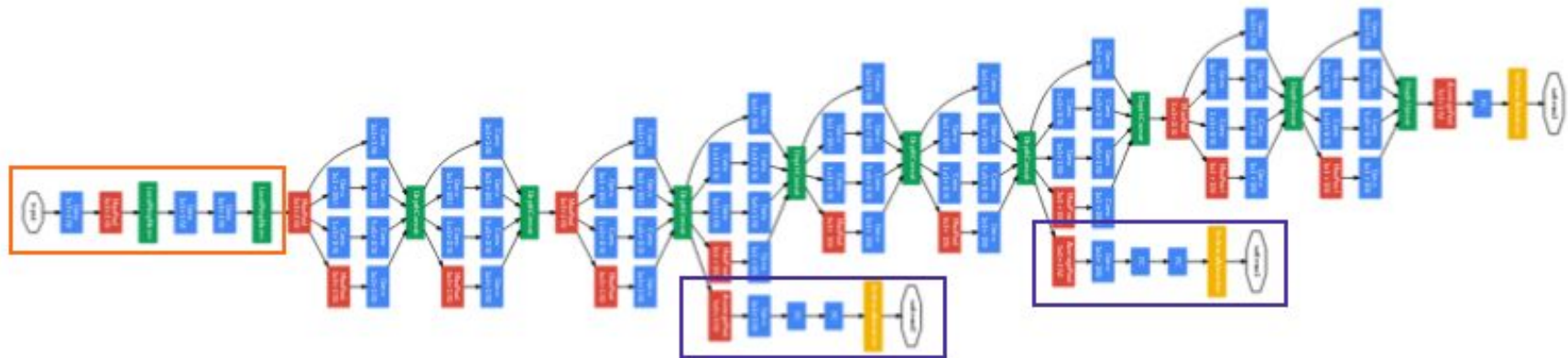


(a) Inception module, naïve version

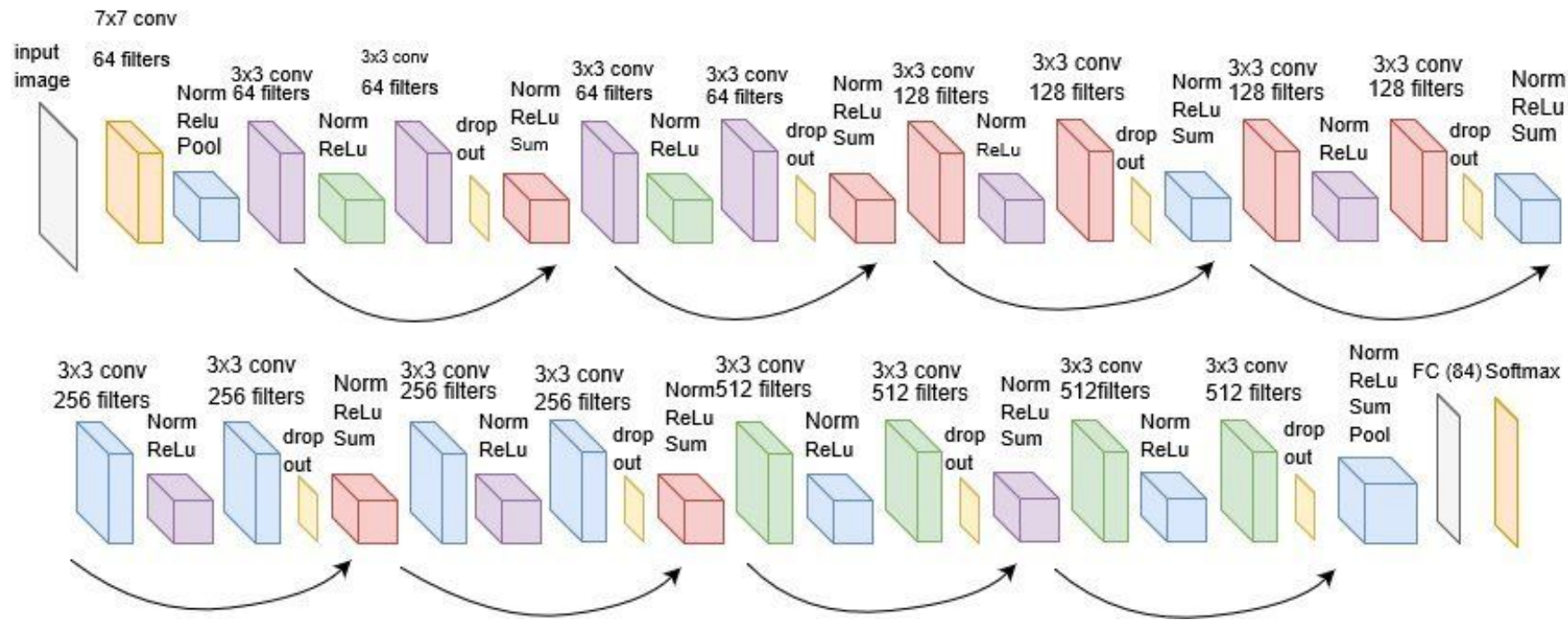


(b) Inception module with dimension reductions

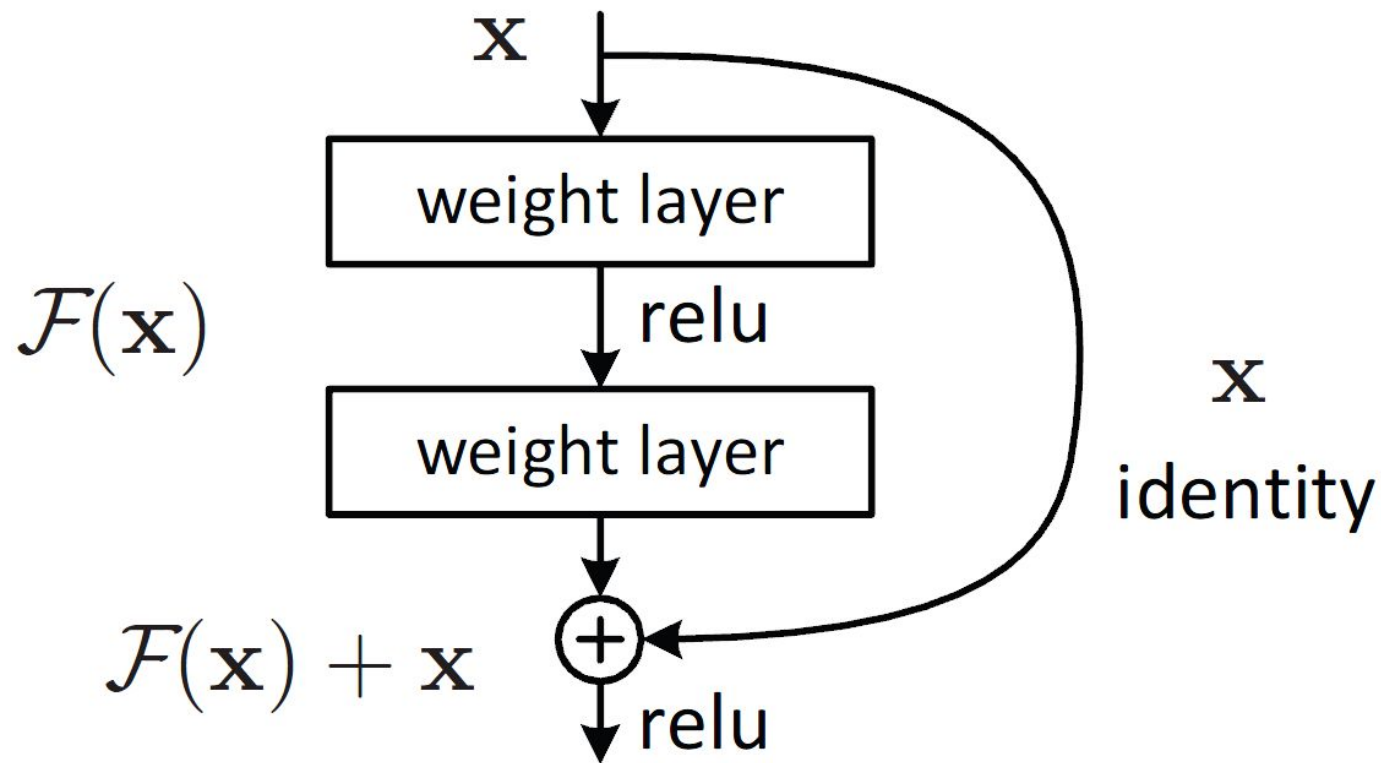
Архитектура Inception-V1



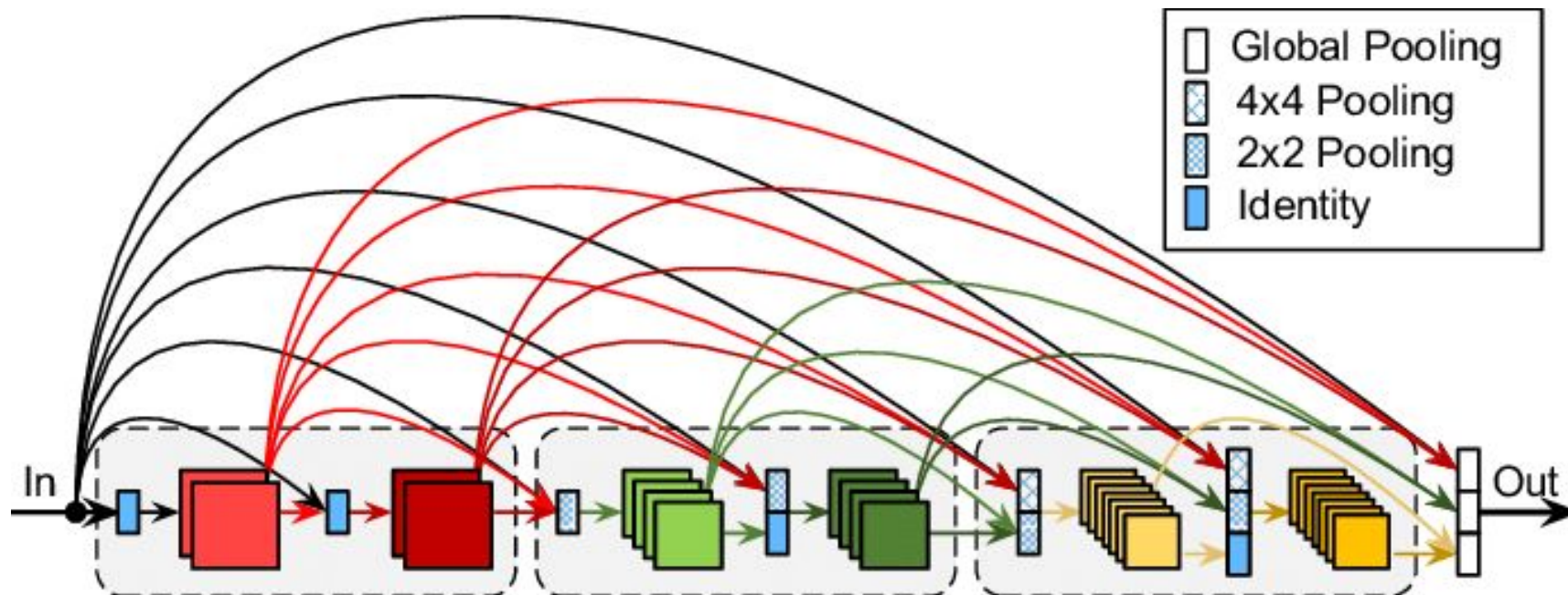
ResNet



Блок ResNet



DenseNet



Вопросы

1. Вход размера $224 \times 224 \times 3$, применяем фильтр размера $11 \times 11 \times 96$ с шагом 4 и дополнением 3. Какого размера будет выход?
2. Зачем делать даунсемплинг?
3. Отличия ResNet от LeNet?

ИСТОЧНИКИ

1. <https://towardsdatascience.com/convolutional-neural-networks-from-the-ground-up-c67bb41454e1>
2. <http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf>
3. <https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>
4. <https://neurohive.io/ru/osnovy-data-science/glubokaya-svernochnaja-nejronnaja-set/>
5. <https://arxiv.org/abs/1512.03385>
6. <https://arxiv.org/pdf/1409.1556>
7. <https://arxiv.org/pdf/1409.4842>