

Свёрточные нейронные сети

Курс «Интеллектуальные информационные системы»

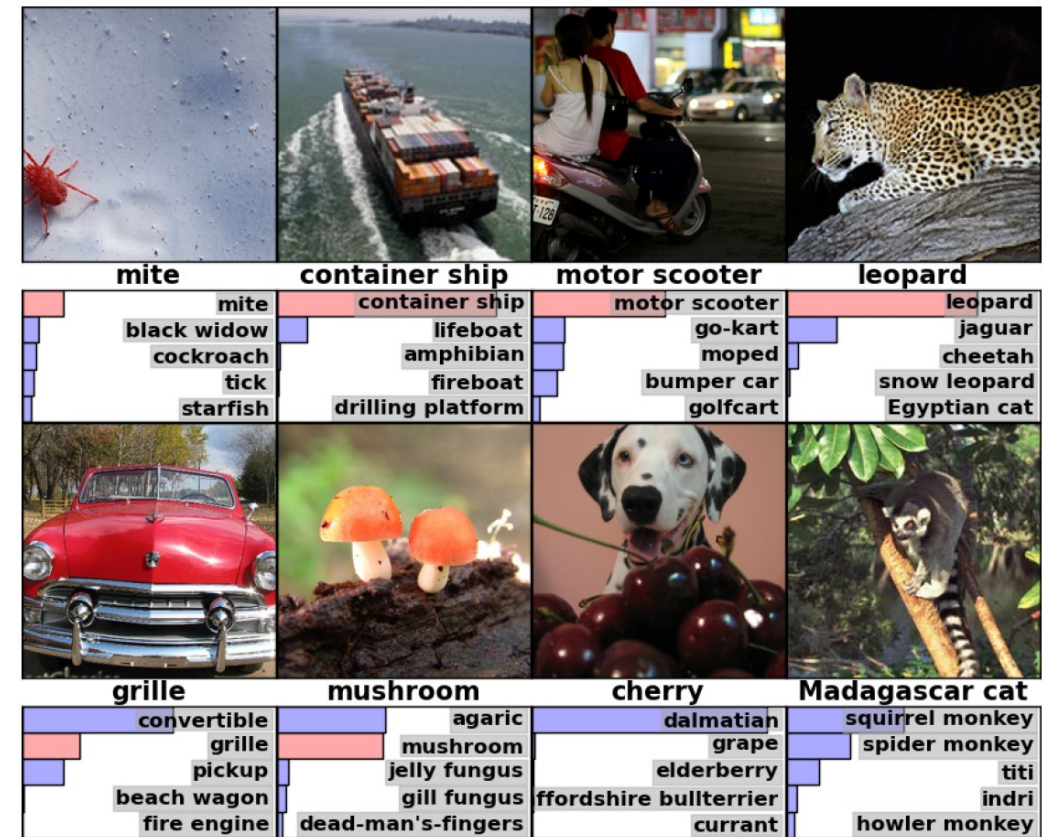
Кафедра управления и информатики НИУ «МЭИ»

Осень 2021 г.

Свёрточные нейронные сети

Convolutional Neural Networks (CNN) - специальная архитектура искусственных нейронных сетей, предложенная Яном Лекуном в 1988 году и нацеленная на эффективное распознавание изображений.

Впервые сверточные нейронные сети (СНС) привлекли всеобщее внимание в 2012 году, когда Алекс Крижевски благодаря им выиграл конкурс ImageNet (грубо говоря, это ежегодная олимпиада по машинному зрению), снизив рекорд ошибок классификации с 26% до 15%, что тогда стало прорывом. На август 2018 года в ImageNet 14 197 122 изображения, разбитых на 21 841 категорию.



<http://neuralnetworksanddeeplearning.com/> - "Нейронные сети и глубинное обучение" Майкла Нильсена

<https://habrahabr.ru/post/309508/> - «Что такое свёрточная нейронная сеть»

Свёрточные нейронные сети



What We See

Интенсивность пикселей

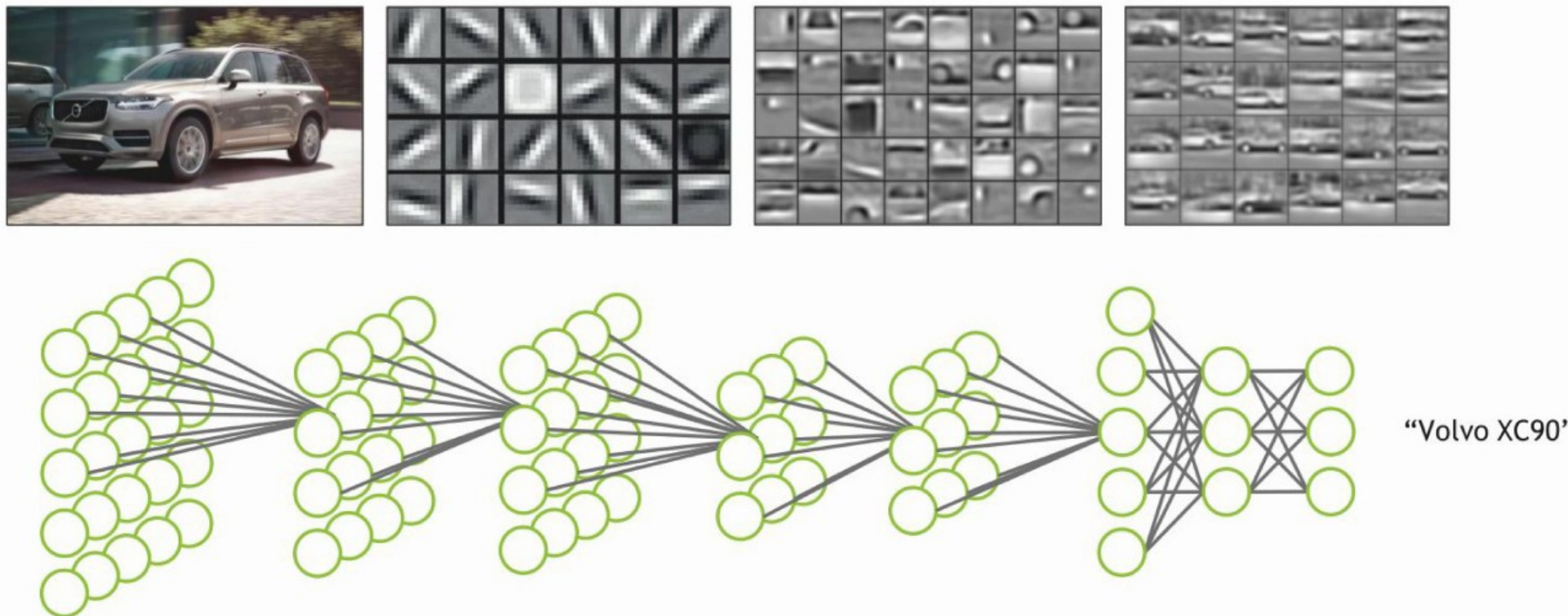
08 02 22 97 38 15 00 40 00 75 04 05 07 78 52 12 50 77 91 08
49 49 99 40 17 81 18 57 60 87 17 40 98 43 69 48 04 56 62 00
81 49 31 73 55 79 14 29 93 71 40 67 53 88 30 03 49 13 36 65
52 70 95 23 04 60 11 42 69 24 68 56 01 32 56 71 37 02 36 91
22 31 16 71 51 67 63 89 41 92 36 54 22 40 40 28 66 33 13 80
24 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 48 38 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 88 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

Свёрточные нейронные сети

СНС в какой то мере - прототип зрительной коры мозга. Они используют некоторые особенности зрительной коры, в которой были открыты так называемые простые клетки, реагирующие на прямые линии под разными углами, и сложные клетки, реакция которых связана с активацией определённого набора простых клеток.

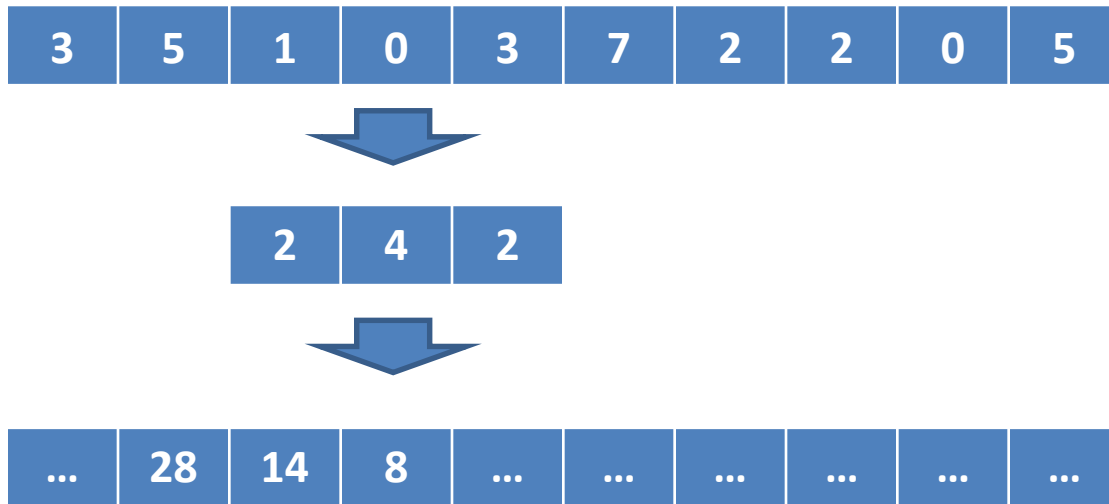
Эту идею детально рассмотрели с помощью потрясающего эксперимента Хьюбел и Визель в 1962 году, в котором показали¹, что отдельные мозговые нервные клетки реагировали (или активировались) только при визуальном восприятии границ определенной ориентации.



¹ <https://www.youtube.com/watch?v=Cw5PKV9Rj3o>

Операция свертки

Название архитектура сети получила из-за наличия операции свёртки, суть которой в том, что каждый фрагмент изображения умножается на матрицу (ядро, фильтр, «окно») свёртки поэлементно, а результат суммируется и записывается в аналогичную позицию выходного изображения. Свертка – способ выделения признаков. Идея в том, чтобы пройти по всему изображению «окном» и посмотреть, насколько данная область изображения похожа на «окно».



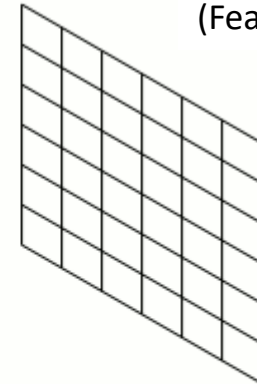
Ядро свертки

0	-1	0
-1	5	-1
0	-1	0

7	6	5	5	6	7
6	4	3	3	4	6
5	3	2	2	3	5
5	3	2	2	3	5
6	4	2	2	3	5
7	6	3	3	4	6
	6	5	5	6	7

Входное
изображение

“Карта признаков”
(Feature map)

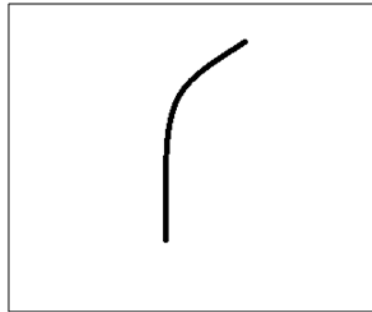


Пример свертки изображения

Каждый фильтр (ядро) можно рассматривать как идентификатор свойства:

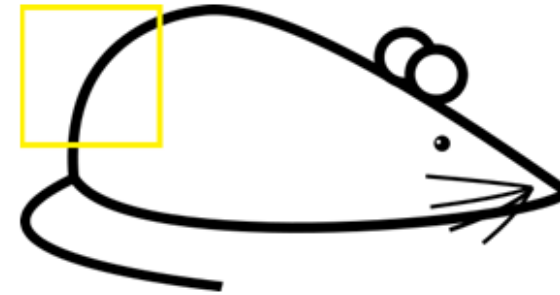
0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter



Visualization of a curve detector filter

Исходное изображение:



Visualization of the filter on the image

1)



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

Пример свертки изображения (2)

2)



Visualization of the filter on the image

0	0	0	0	0	0	0
0	40	0	0	0	0	0
40	0	40	0	0	0	0
40	20	0	0	0	0	0
0	50	0	0	0	0	0
0	0	50	0	0	0	0
25	25	0	50	0	0	0

Pixel representation of 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 = 0

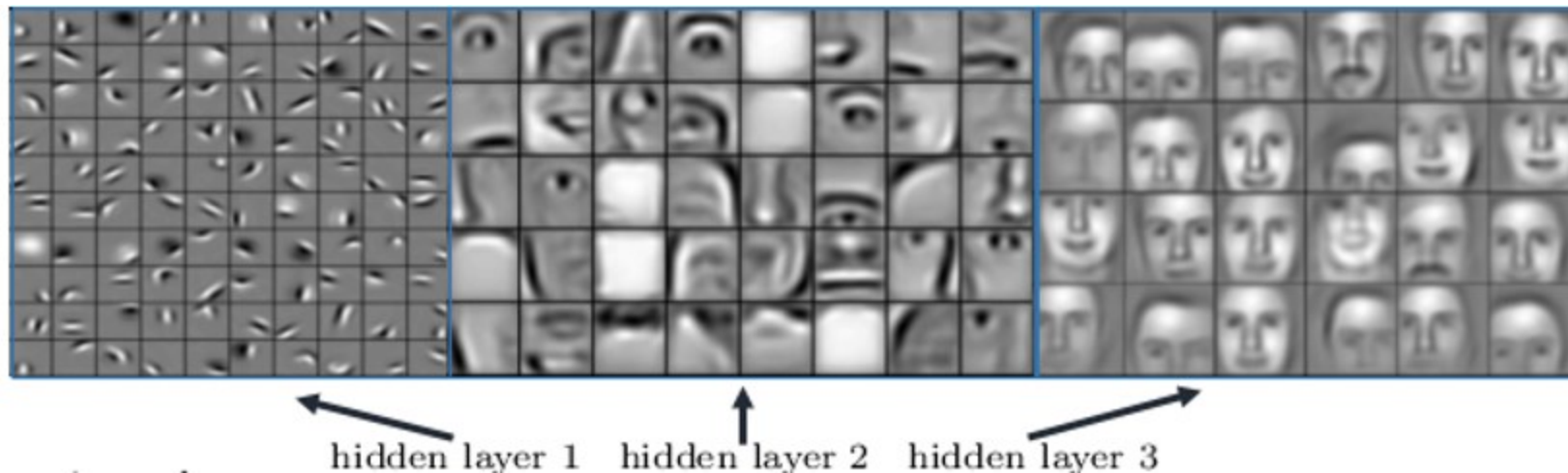
В первом случае большое значение после свертки говорит, что в данной области присутствует что-то похожее на кривую.

Во втором случае – похожей кривой не было.

Данный результат получен только для одного фильтра. Это фильтр, который обнаруживает линии с изгибом наружу. Могут быть другие фильтры для линий, изогнутых внутрь или просто прямых. Чем больше фильтров, тем больше глубина карты свойств, и тем больше информации мы имеем о вводной картинке.

Сверточные нейронные сети

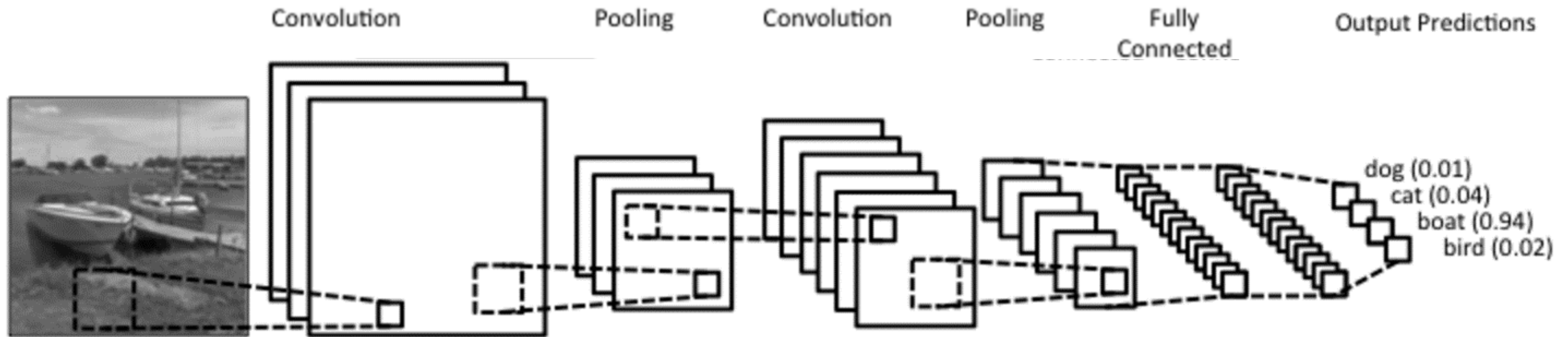
Примеры фактических визуализаций фильтров нескольких свёрточных слоев обученной сети:



Чем дальше мы движемся вглубь нейросети, тем более сложные характеристики объектов будут выделяться: кольца, квадраты, пересечения линий. В конце сети могут быть фильтры, которые активируются при наличии лица, рукописного текста, автомобиля итд.

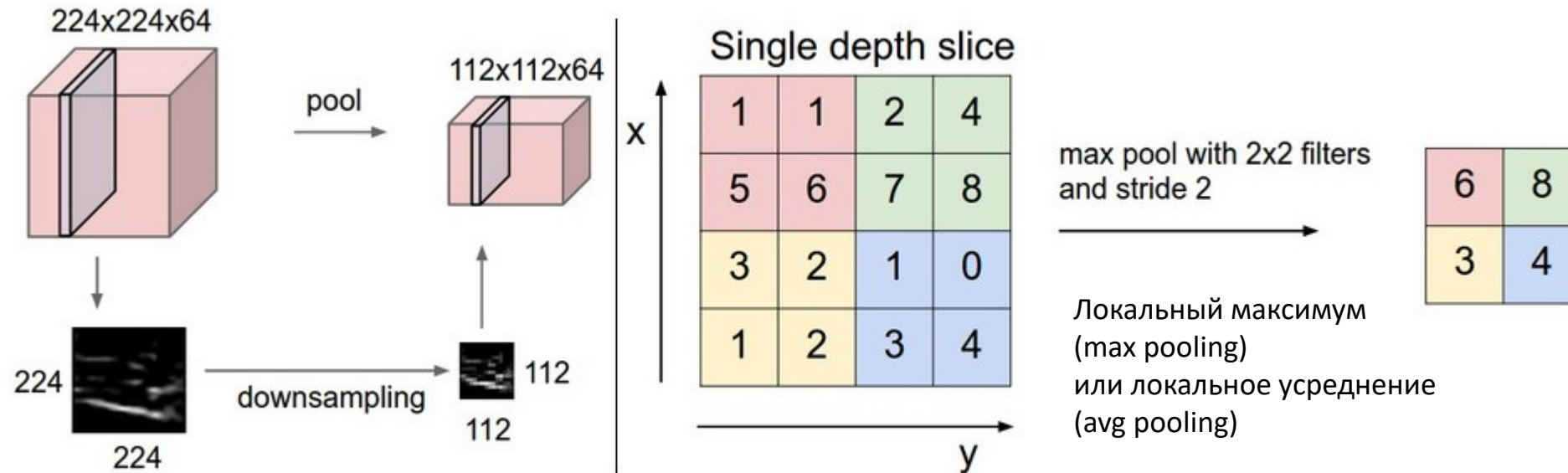
<https://www.youtube.com/watch?v=AgkfIQ4IGaM>

Схема сверточной нейронной сети



- Input – входное изображение (например, $128 \times 128 \times 1$, где 128 — ширина и высота изображения, 1 — один цветовой канал)
- Convolutional – сверточный слой (если используется 3 фильтра, то объем будет равен $128 \times 128 \times 3$)
- Pooling или Subsampling – слой пулинга (подвыборки или субдискретизации). Размер уменьшается до, например, $64 \times 64 \times 3$.
- Fully Connected - полносвязный слой - выводит N-мерный вектор (N — число классов) для определения нужного класса (обычная нейросеть)

Пулинг



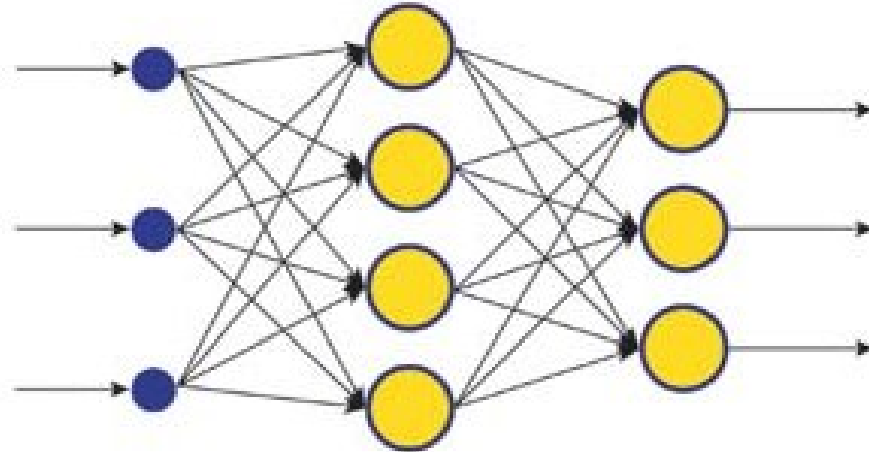
Использование этого слоя позволяет улучшить распознавание образцов с изменённым масштабом (уменьшенных или увеличенных).

Логика работы такова: если на предыдущей операции свертки уже были выявлены некоторые признаки, то для дальнейшей обработки настолько подробное изображение уже не нужно, и оно уплотняется до менее подробной картинке.

Полносвязный слой

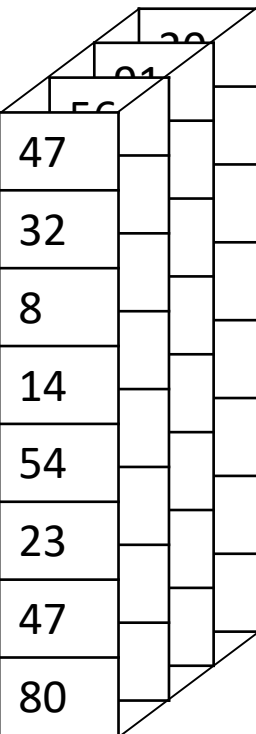
Теперь, когда мы можем обнаружить высокоуровневые свойства, мы прикрепляем полносвязный слой в конце сети.

Этот слой берёт вводные данные и выводит N-мерный вектор с вероятностями каждого из N классов.



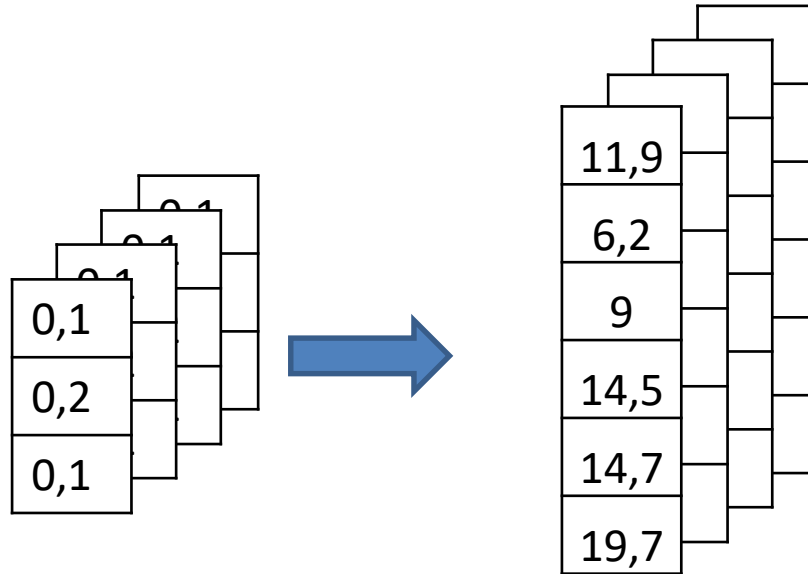
Одномерные свёрточные НС

Для анализа текстовых данных применяются одномерные свёрточные нейронные сети. Слово представляется в векторном виде:



the	47
quick	32
brown	8
fox	14
jumps	54
over	23
the	47
lazy	80

Текст в векторном виде



Ядро свертки

Карта признаков