

Нейронные сети и их практическое применение.

Лекция 9. Сверточные сети.

Дмитрий Буряк
к.ф.-м.н
dyb04@yandex.ru

Глубокое обучение (Deep learning)

❑ Общие сведения

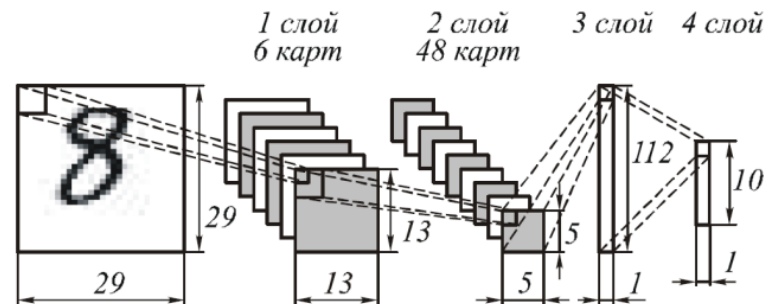
- тип алгоритмов машинного обучения, которые пытаются моделировать высокоуровневые абстракции в данных, используя архитектуры, состоящие из множества нелинейных трансформаций;
- большое число слоев и нейронов -> большое число весов;
- высокие требования к вычислительным ресурсам.

❑ Типы архитектур

- многослойные сети
- сверточные нейронные сети
- глубокие сети доверия

❑ Решаемые задачи

- компьютерное зрение
- распознавание речи
- обработка естественного языка



Свертка

17	24	1	8	15
23	5	7	14	16
4	6	13	20	22
10	12	19	21	3
11	18	25	2	9

X

8	1	6
3	5	7
4	9	2

→

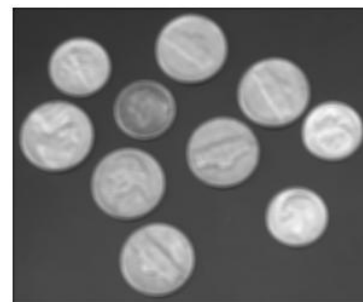
17	24	1*2	8*9	15*4
23	5	7*7	14*5	16*3
4	6	13*6	20*1	22*8
10	12	19	21	3
11	18	25	2	9



X

0.04	0.04	0.04	0.04	0.04
0.04	0.04	0.04	0.04	0.04
0.04	0.04	0.04	0.04	0.04
0.04	0.04	0.04	0.04	0.04
0.04	0.04	0.04	0.04	0.04

=



X

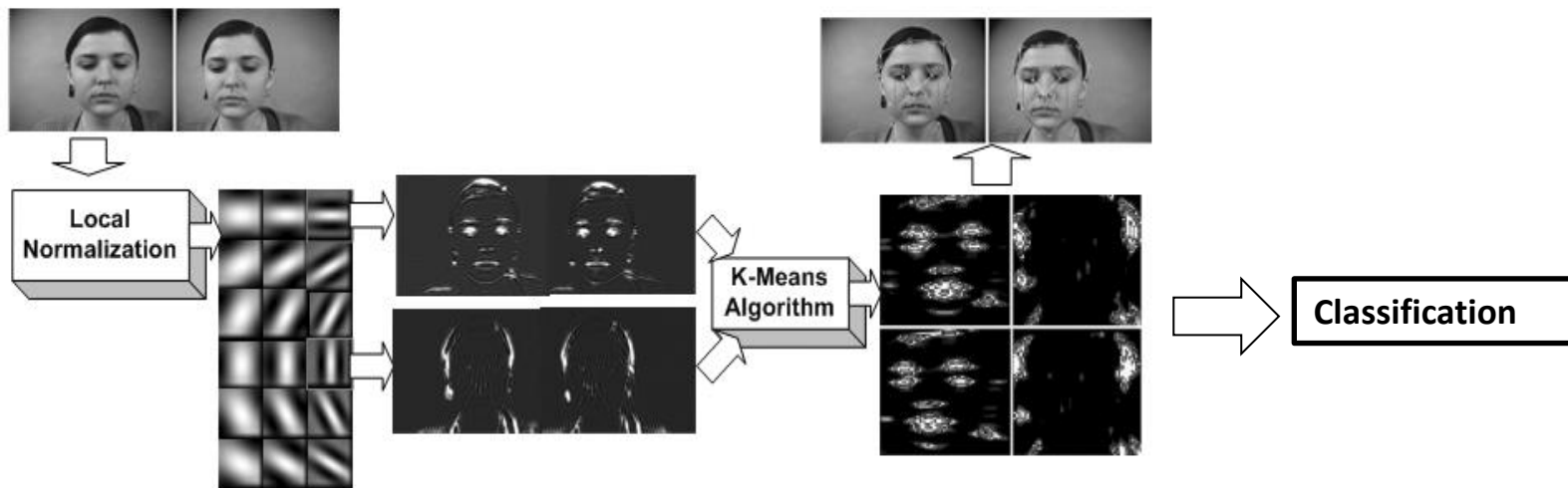
-0,17	-0,67	-0,17
-0,67	4,36	-0,67
-0,17	-0,67	-0,17

=



Применение свертки для задач распознавания

- ❑ Решение задачи обнаружения лица (Bazanov et al., 2004)

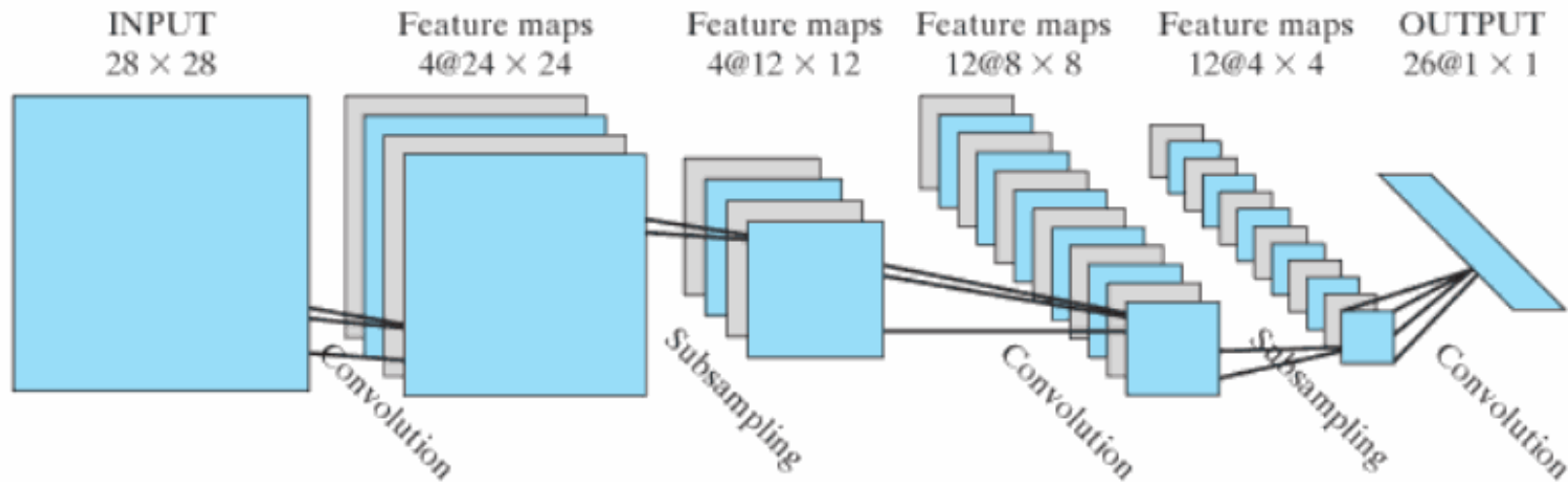


- ❑ Сверточные НС – поиск оптимального набора операций свертки

Сверточные сети

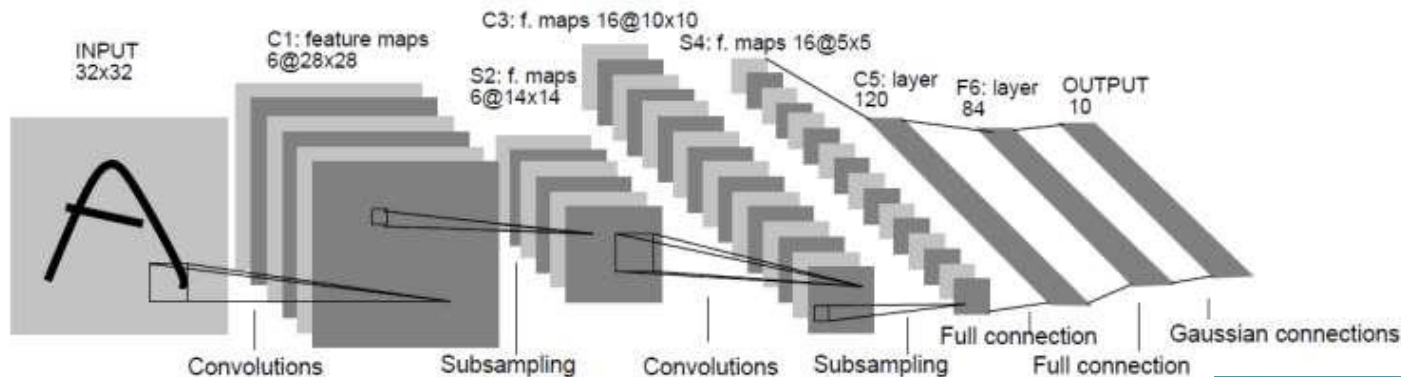
- Многослойный переспетрон для распознавания 2D форм
- Разделяемые веса
- Устойчивость к сдвигу, повороту, изменению размера и другим искажениям
- Алгоритм обучения: обратное распространение ошибки
- Автоматическое выделение особенностей изображения

Сверточные сети: архитектура

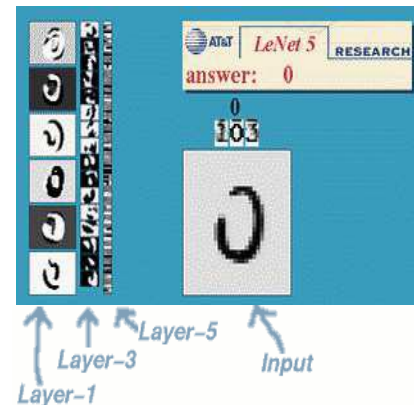


- Свертка
- Усреднение
- Общие веса

Сверточные сети: распознавание СИМВОЛОВ



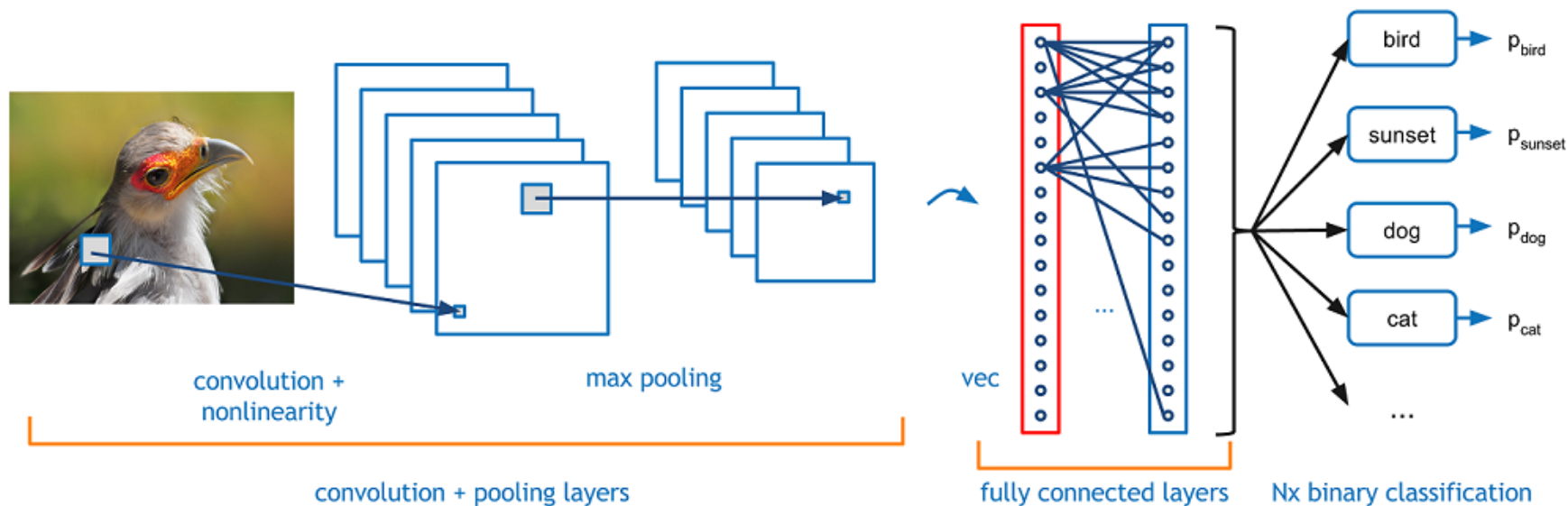
- Обработка локальных областей
- Совместное использование весов:
100000 связей, 2800 настраиваемых весов
- Высокая обобщающая способность
- Возможность распараллеливания
- Устойчивость к искажениям и шуму



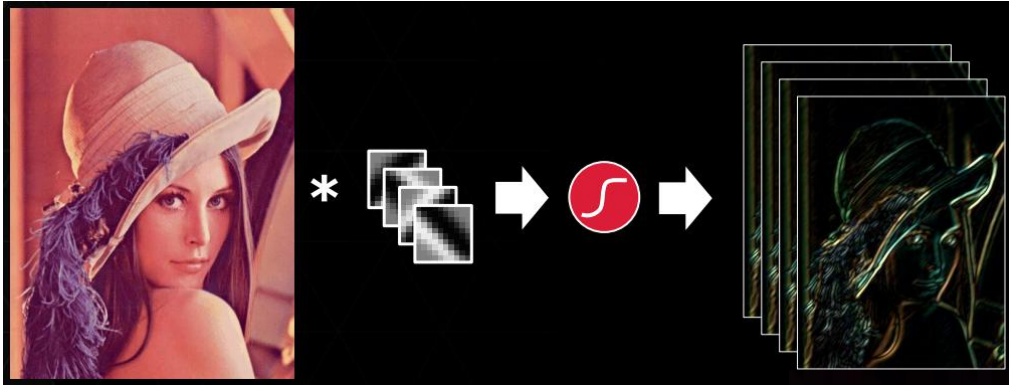
Сверточные сети: результаты

Сеть	Процент ошибок на тестовой выборке (%)
Линейный классификатор (1-слойная НС)	12.0
2-слойная НС, 300 скрытых	4.7
2- слойная НС, 1000 скрытых элементов	4.5
3- слойная НС, 300+100 скрытых элементов	3.05
3- слойная НС, 500+150 скрытых элементов	2.95
Большая сверточная сеть	0.39

Типы слоев



Типы слоев. Свертка



- Свертка входного образа с набором ядер
- Применение нелинейной функции активации
- Результат – множество карт признаков

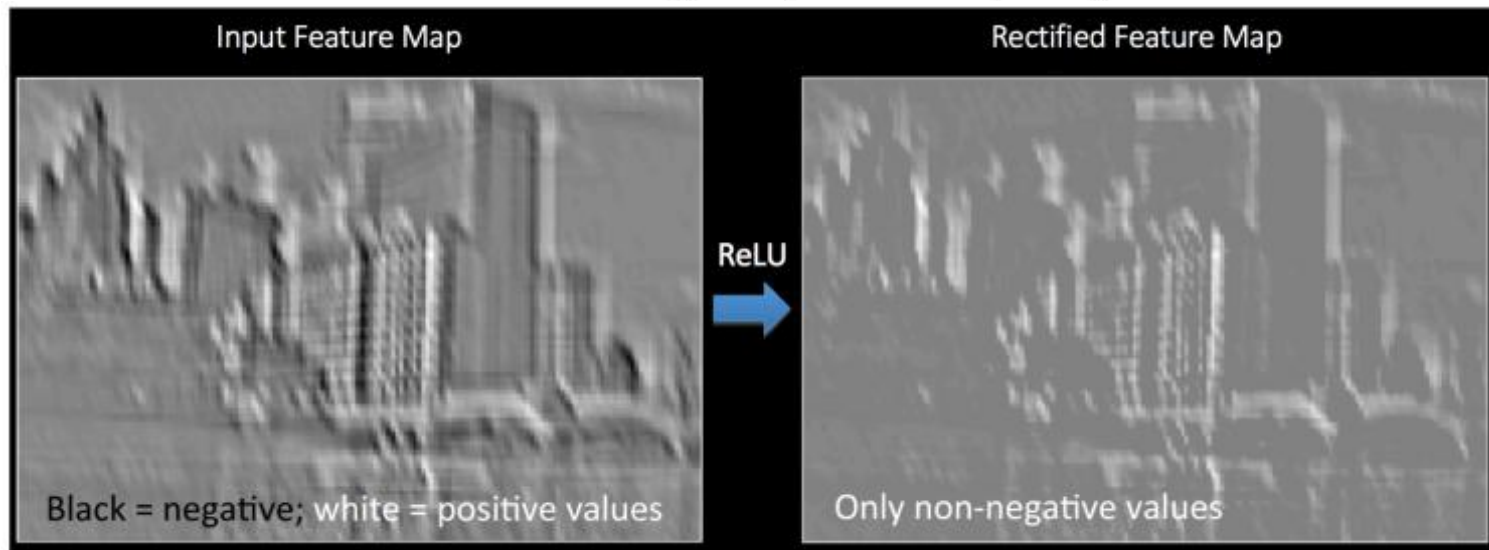
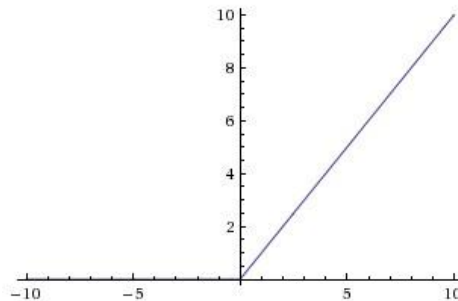


Input

Типы слоев. Функция активации

□ ReLU – Rectified Linear Unit:

$$f(u) = \max(0, u)$$



Типы слоев. Свойства ReLu

❑ Преимущества

- Решение проблемы «vanishing gradient»
- Проще для вычисления

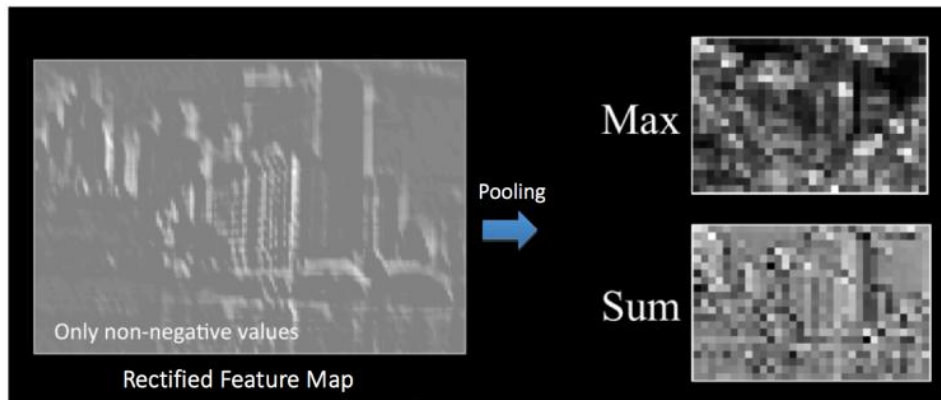
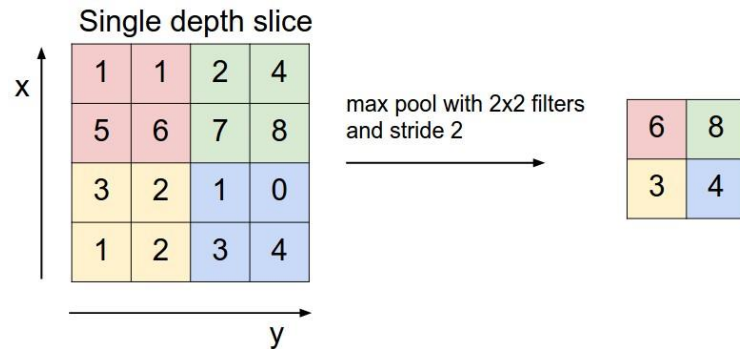
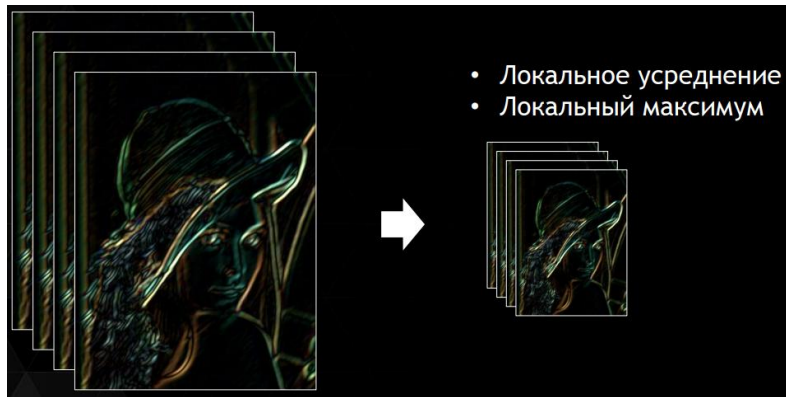
❑ Недостатки

- Функция не ограничена
- Возможность переобучения

❑ ReLu vs сигмоида

- Для ряда задач важны ограниченность и симметричность
- В 95% случаев ReLu эффективнее

Типы слоев. Понижение размерности (pooling)



Типы слоев. Полносвязанный слой

- Архитектура аналогичная персептрону
- Один или несколько слоев
- Проведение классификации на основе выделенных признаков

Обучение

❑ Основные проблемы

- Большое число связей
- Размер обучающей выборки
- Переобучение
- Затухание градиента

❑ Методы

- Обратное распространение ошибки
- Стохастический градиентный спуск
- Регуляризация
- Dropout
- Batch normalization

База изображений ImageNet

❑ ImageNet

- Создана в 2009
- Более 14М изображений с аннотациями
- 1М изображений с отмеченными объектами
- Более 20К классов (несколько сотен изображений каждого класса)



❑ ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

- Проводится с 2010г
- Классификация изображений, обнаружение объектов



❑ ILSVRC. Классификация изображений

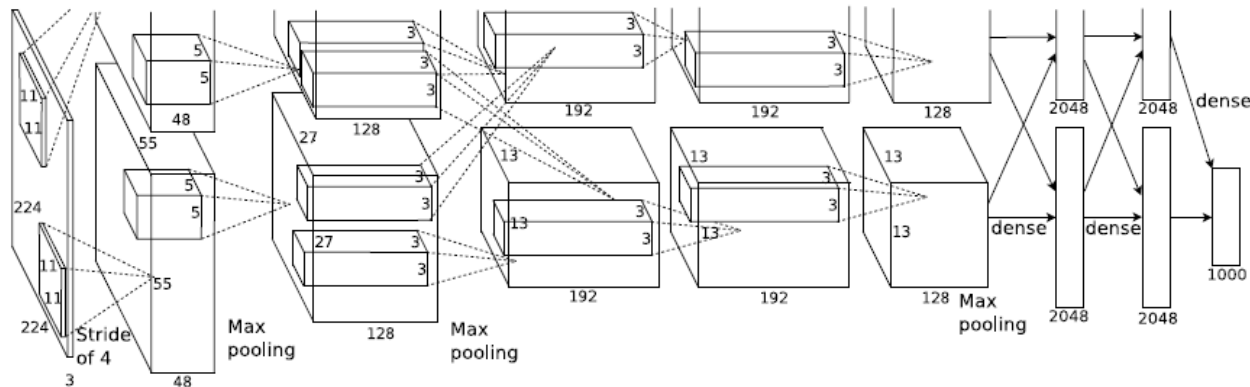
- 1.2М изображений (обучающая выборка)
- 100К тестовых изображений
- 1000 классов



Победитель ILSVRC-2012. AlexNet

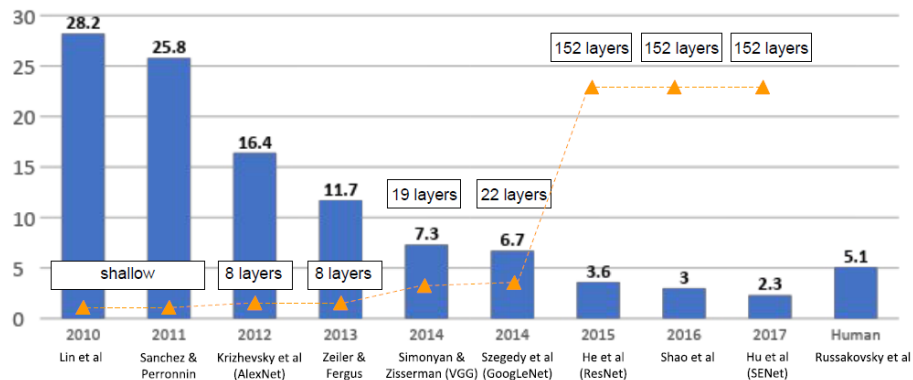
□ AlexNet (A. Krizhevsky et al, 2012)

- Первая сверточная сеть победитель ILSVRC;
- Ошибка Top-5: **15.3%** (второй результат SIFT – 26.2%);
- Сверточная НС, 8 слоев (5 сверточных), 60М параметров;
- ReLU, Dropout, Увеличение обучающей выборки (сдвиг, отражение);
- Обучение заняло 5 дней на нескольких GPU.

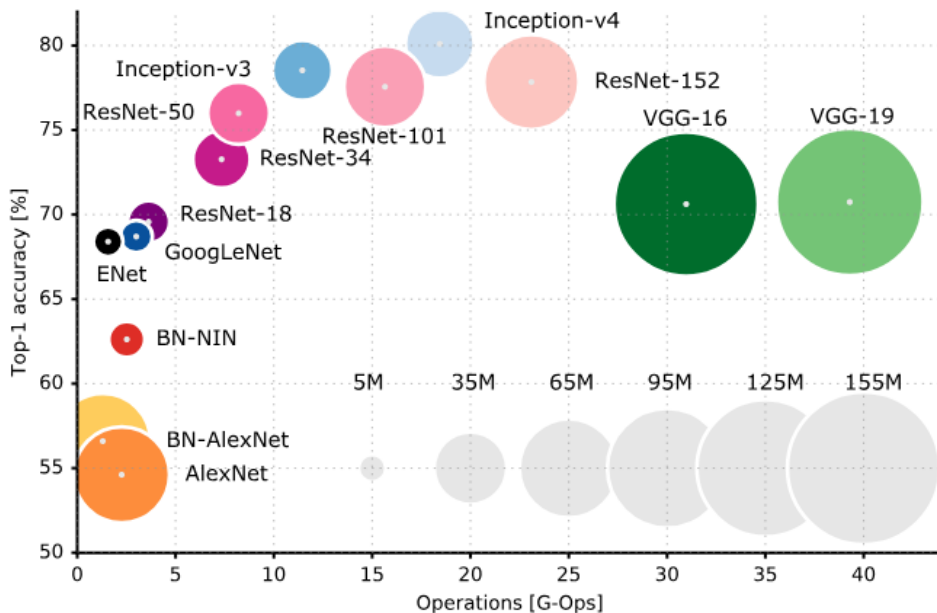


Развитие НС для ILSVRC

ImageNet Large Scale Visual Recognition Challenge (ILSVRC) winners



http://cs231n.stanford.edu/slides/2018/cs231n_2018_lecture09.pdf



Классические архитектуры

❑ VGG (K.Simonyan, A. Zisserman, 2014) : **7.3%** (Top-5)

- Классическая сверточная архитектура (ядра 3x3);
- 138M параметров, 16 слоев.

❑ GoogLeNet (K. Szegedy et al, 2015): **6.7%** (Top-5)

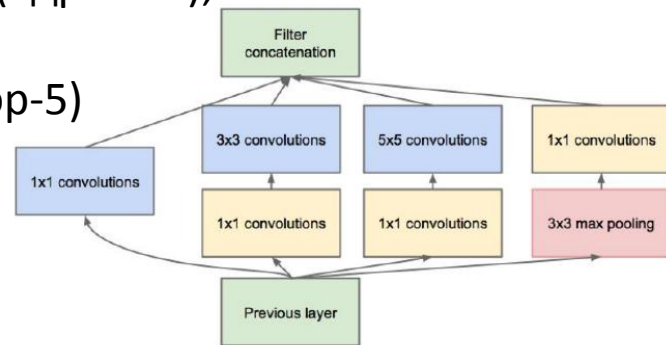
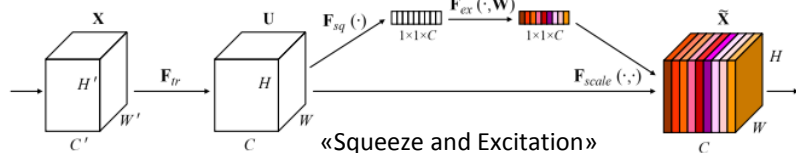
- Inception блок
- 60M параметров, 50 слоев

❑ ResNet (K. He et al, 2015): **3.57%** (Top-5)

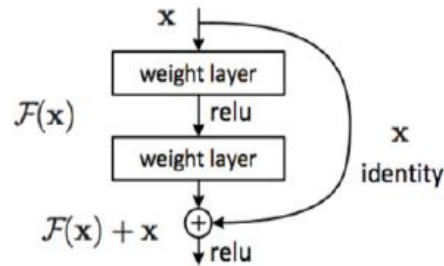
- Residual connections
- 60M параметров, 152 слоя

❑ SENet-154 (J. Hu et al, 2017): **2.25%** (Top-5)

- «Squeeze and Excitation» (SE) модуль
- 60M параметров, 154 слоя



GoogLeNet Inception блок

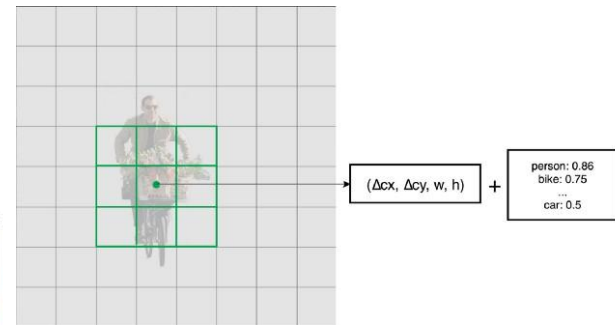
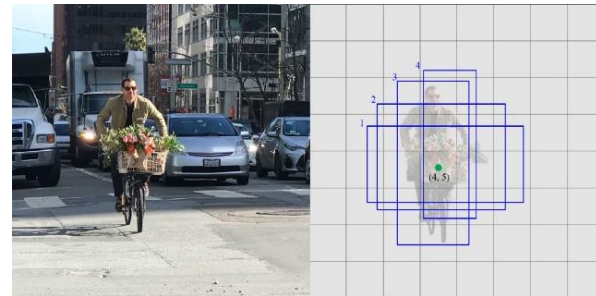


Residual connection блок

D
16 weight layers
conv3-64 conv3-64
conv3-128 conv3-128
conv3-256 conv3-256 conv3-256
conv3-512 conv3-512 conv3-512
conv3-512 conv3-512 conv3-512
maxpool FC-4096 FC-4096 FC-1000 soft-max

Примеры. Обнаружение объектов

- ❑ SSD: Single Shot Detector (W. Liu et al., 2016);
- ❑ Оценка «формы» объекта для каждой ячейки выходной карты;
- ❑ Оценка положения и вероятности класса;
- ❑ Обнаружение на разных масштабах;
- ❑ Non-maximum suppression;
- ❑ **mAP = 83.2% (PASCAL VOC 2007)**
- ❑ **mAP = 82.2% (PASCAL VOC 2012)**



Вопросы

1. Какой метод обработки данных заложен в архитектуру сверточной сети?
2. Как функционирует слой свертки?
3. Как работает слой понижения размерности?