

# Основы практического использования нейронных сетей.

Лекция 8. Обзор современных глубоких сетей для  
классификации изображений.

Дмитрий Буряк.

к.ф.-м.н.

[dyb04@yandex.ru](mailto:dyb04@yandex.ru)

# Задача классификации изображений

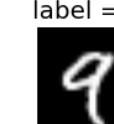
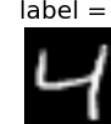
- Множество изображений.
- Каждое изображение содержит один объект какого-то класса.
- Классы объектов заданы.
- Необходимо определить к какому из известных классов относится изображенный объект.
- Не требуется определять позицию, где находится объект.
- Два варианта задачи классификации:
  - Каждое изображение содержит только объекты заданных классов;
  - Есть изображения объектов не из заданного набора классов.

Точность алгоритма классификации:  $\frac{M}{N} * 100\% ,$

М – число изображений с правильно определенным классом,

Н – общее количество изображений.

- Пример: классификация изображений MNIST



# 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)

- Проводился с 2010г по 2017г;
- Классификация изображений, обнаружение объектов.



## □ ILSVRC. Классификация изображений

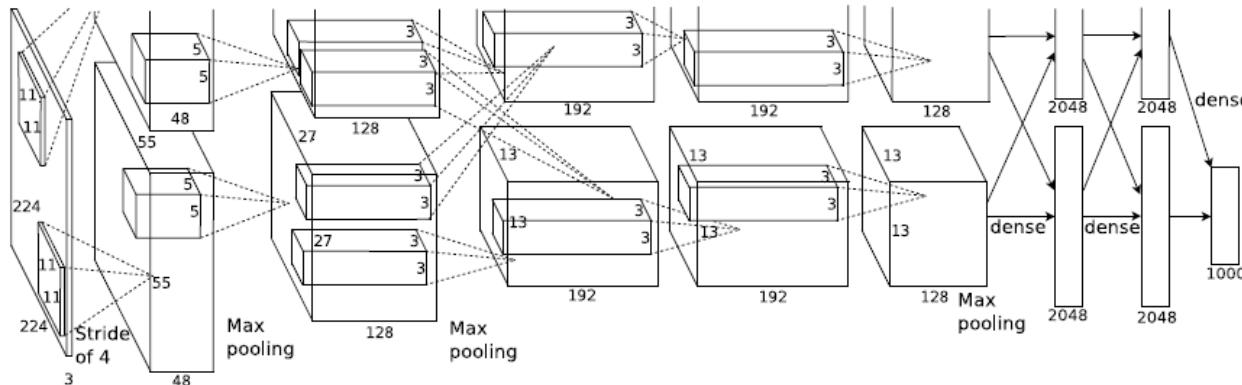
- 1.2М изображений (обучающая выборка)
- 100K тестовых изображений
- 1000 классов



# Победитель ILSVRC-2012. AlexNet

## □ AlexNet (A. Krizhevsky et al, 2012)

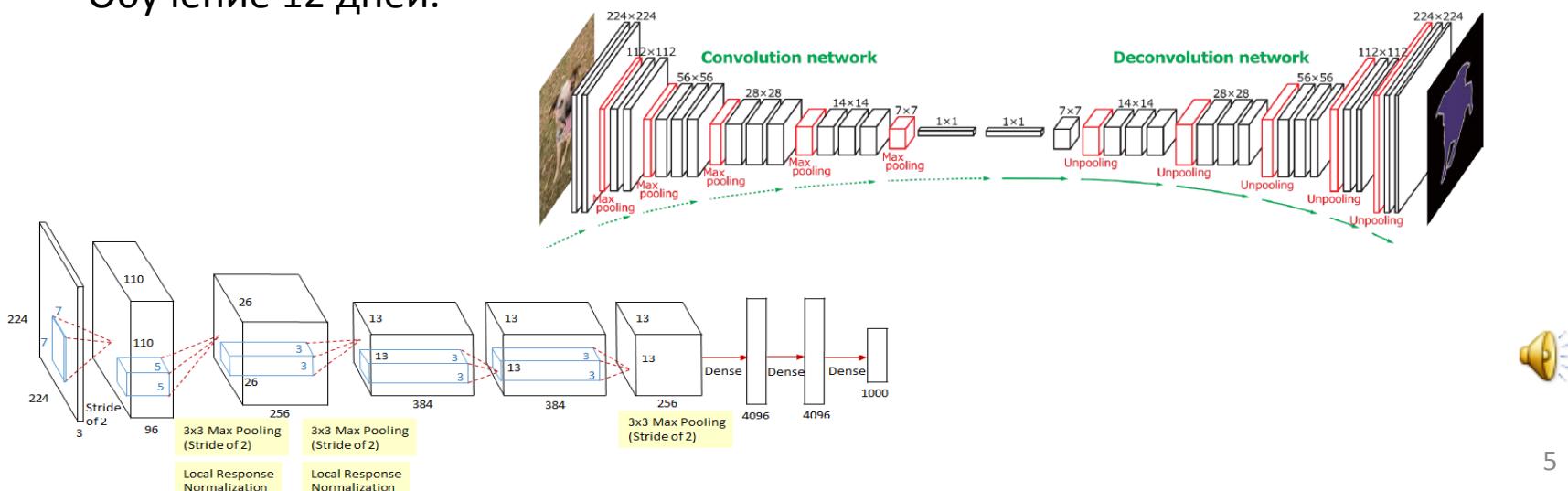
- Первая сверточная сеть победитель ILSVRC;
- Ошибка Top-5: **15.3%** (второй результат SIFT – 26.2%);
- Сверточная НС, 8 слоев (5 сверточных), 60М параметров;
- ReLU, Dropout, Увеличение обучающей выборки (сдвиг, отражение);
- Обучение заняло 5 дней на нескольких GPU.



# Победитель ILSVRC-2013. Clarifai

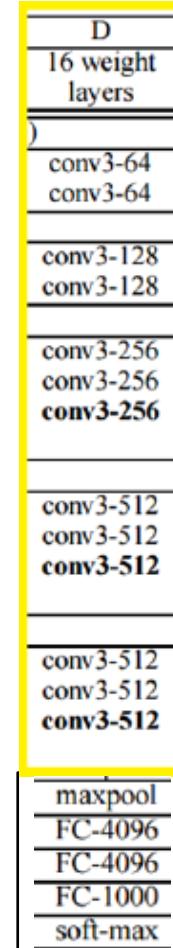
□ ZFNet (Clarifai – стартап авторов ZFNet) – (M. Zeiler et al, 2014)

- Ошибка Top-5: **11.2%** (несколько сетей);
- Визуализация активаций внутренних слоев - Deconvnet;
- Анализ внутренних шаблонов → тонкая настройка AlexNet;
- Фильтры первого слоя  $(11 \times 11) \rightarrow (7 \times 7)$ ; шаг свертки первого слоя  $4 \rightarrow 2$ ;
- Обучение 12 дней.



# VGG Net

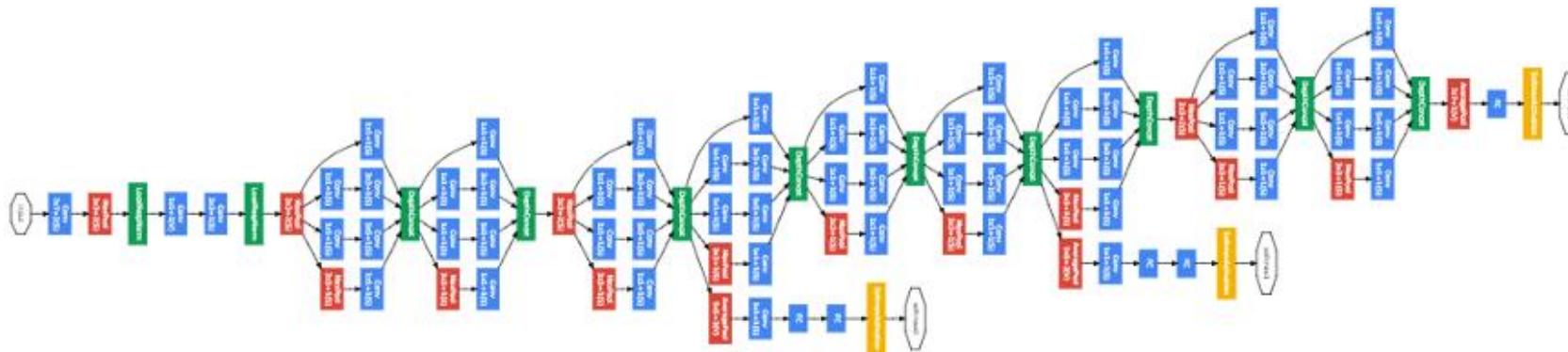
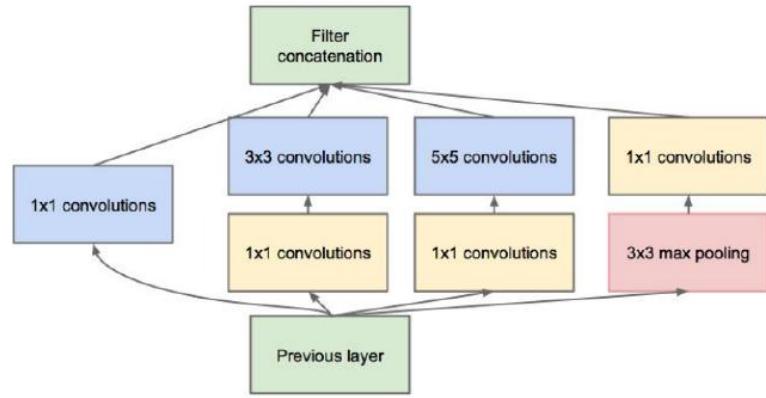
- VGG (K.Simonyan, A. Zisserman, 2014)
  - Ошибка Top-5: **7.3%** (ILSVRC-2014);
  - Классическая CNN;
  - Применение ядер 3x3;
  - $(3 \times 3) + (3 \times 3) \sim (5 \times 5)$ ;  $(3 \times 3) + (3 \times 3) + (3 \times 3) \sim (7 \times 7)$ ;
  - Обучение 4GPUx(3 недели);
  - Активно применяется при решении задач.



# Победитель ILSVRC-2014. GoogLeNet

## □ GoogLeNet – (K. Szegedy et al, 2015)

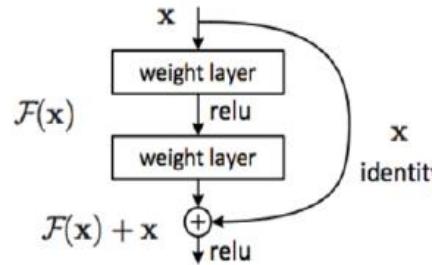
- Ошибка Top-5: **6.7%**;
- «Параллельные свертки-Конкатенация»: Inception модуль.
- 22 модуля, более 50 сверточных слоев;
- Несколько выходных слоев;
- Размер 55Мб (VGG16 – 490Мб);
- Обучение 7 дней на нескольких GPU;



# Победитель ILSVRC-2015. ResNet

## □ ResNet – (K. He et al, 2015)

- Ошибка Top-5: **3.57%**;
- Рост глубины → меньше ошибки → сложность оптимизации;
- «Остаточные» связи – единичное отображение, нет дополнительных параметров;
- «Остаточные» соединения → лучше распространение градиента
- 152 слоя, ядра 3x3;
- Обучение 21 день на 8 GPU.

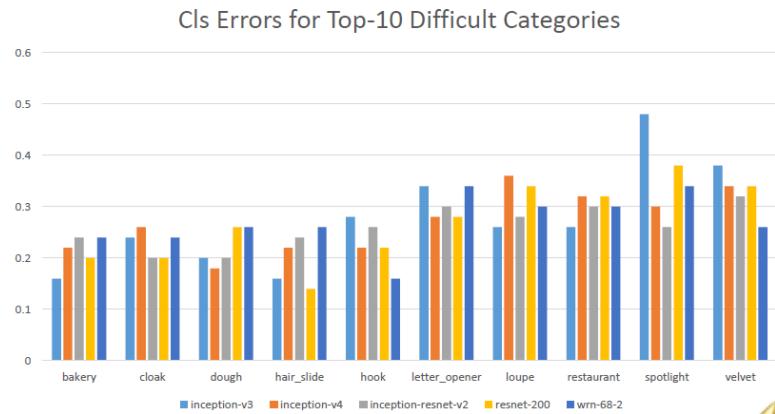


# Победитель ILSVRC-2016. Trimpson-Soushen

## □ Trimpson-Soushen

- Ошибка Top-5: **2.99%**;
- Использование предобученных сетей для выявления Топ-10 наиболее сложных категорий в ImageNet;
- Inception (v3, v4), Inception-ResNet-v2, Pre-Activation ResNet-200, Wide ResNet;
- Нет лучшей сети для всех категорий из Топ-10 + слабая коррелированность решений → объединение решений.

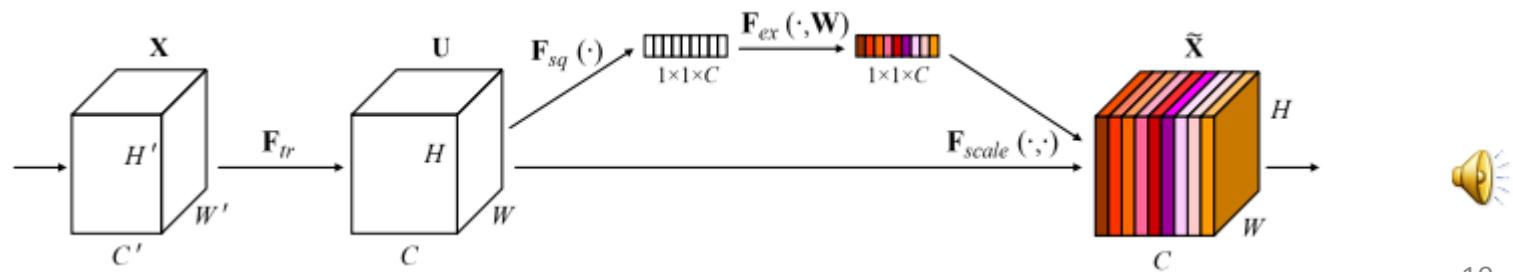
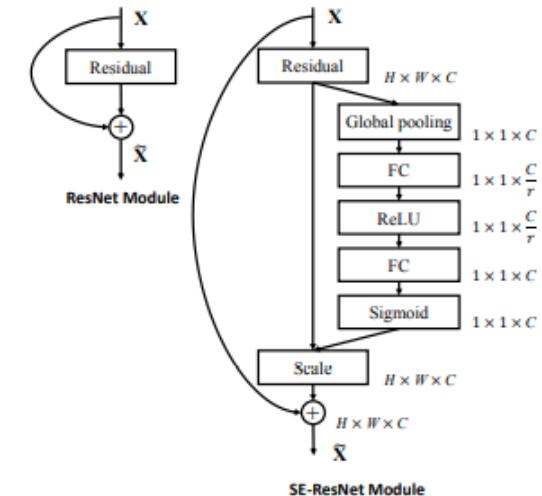
	Inception-v3	Inception-v4	Inception-Resnet-v2	Resnet-200	Wrn-68-3	Fusion (Val.)	Fusion (Test)
Err. (%)	4.20	4.01	3.52	4.26	4.65	2.92 (-0.6)	2.99



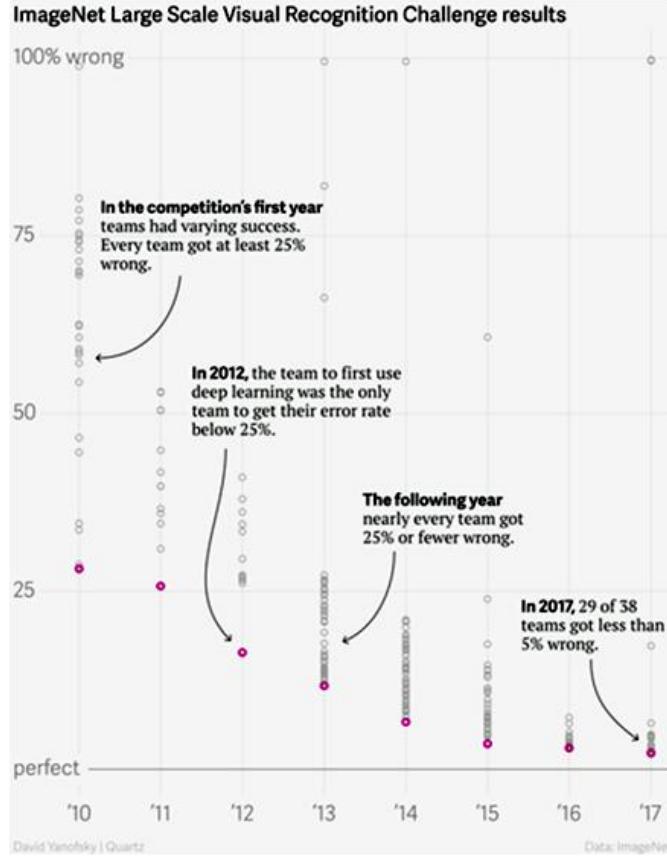
# Победитель ILSVRC-2017. SENet

## □ SENet-154 – (J. Hu et al, 2017)

- Ошибка Top-5: **2.25%**;
- Взвешивание карт признаков;
- «Squeeze and Excitation» (SE) модуль;
- Интеграция в различные существующие архитектуры: Inception, ResNet и т.д.
- Победитель SE+ResNeXt;
- Обучение на 64 GPU для поддержки пакетов размером 2048 изображений.



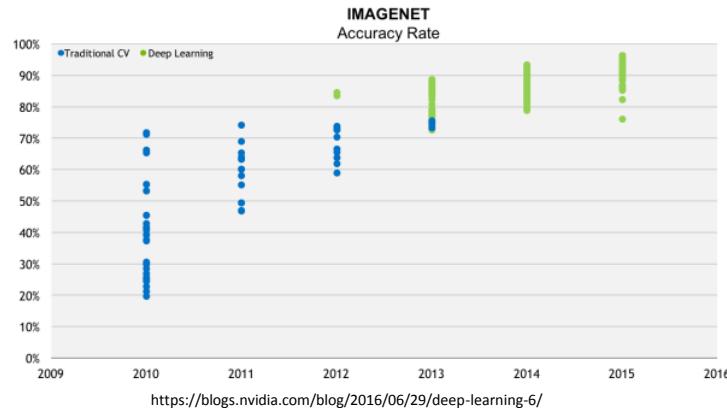
# Статистика ILSVRC



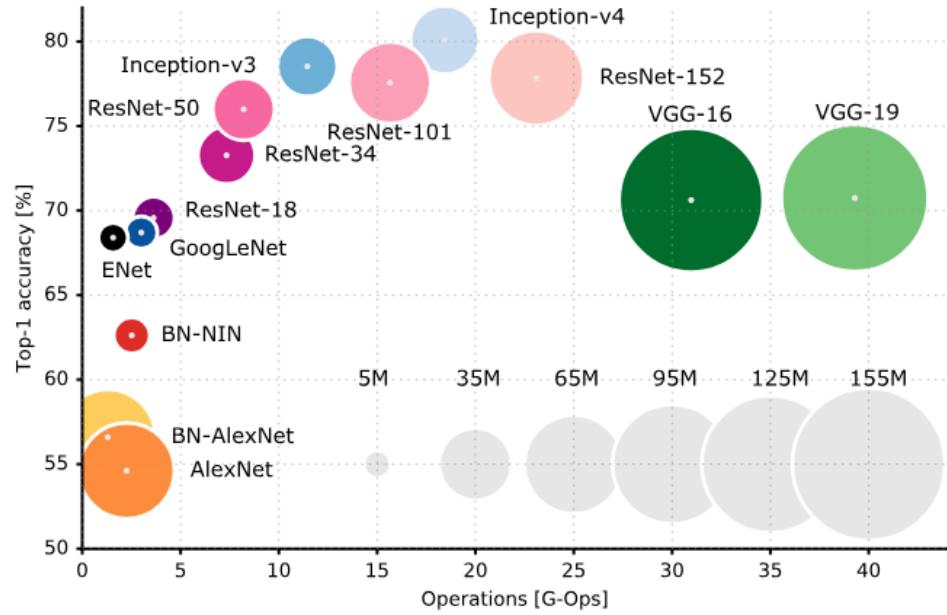
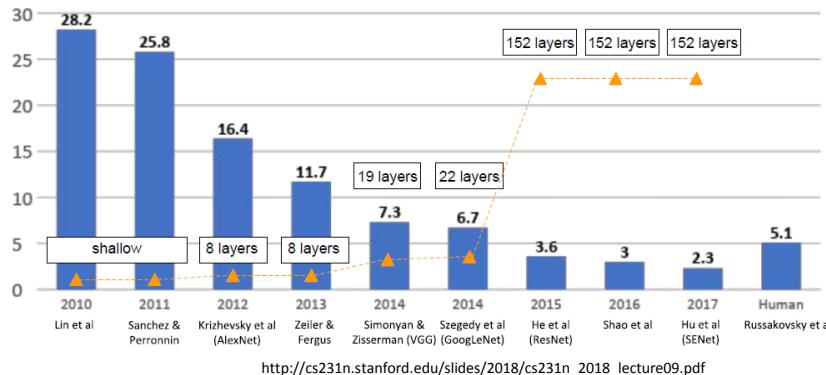
Сеть	Ошибка Top-5
AlexNet	15.3%
ZFNet	11.2%
VGG	7.3%
GoogLeNet	6.7%
ResNet	3.57%
Trimpson-Soushen	2.99%
SENet-154	2.25%



# Развитие НС для ILSVRC



ImageNet Large Scale Visual Recognition Challenge (ILSVRC) winners



# Ссылки

- ❑ Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.
- ❑ Zeiler, Matthew D., and Rob Fergus. "Visualizing and understanding convolutional networks." *European Conference on Computer Vision*. Springer International Publishing, 2014.
- ❑ Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for largescale image recognition." *arXiv preprint arXiv:1409.1556*(2014).
- ❑ Szegedy, Christian, et al. "Going deeper with convolutions." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015.
- ❑ He, Kaiming, et al. "Deep residual learning for image recognition." *arXiv preprint arXiv:1512.03385* (2015).
- ❑ Hu, Jie, et al. "Squeeze-and-Excitation Networks." *arXiv preprint arXiv:1709.01507* (2017).