

# Практический анализ данных и машинное обучение: искусственные нейронные сети

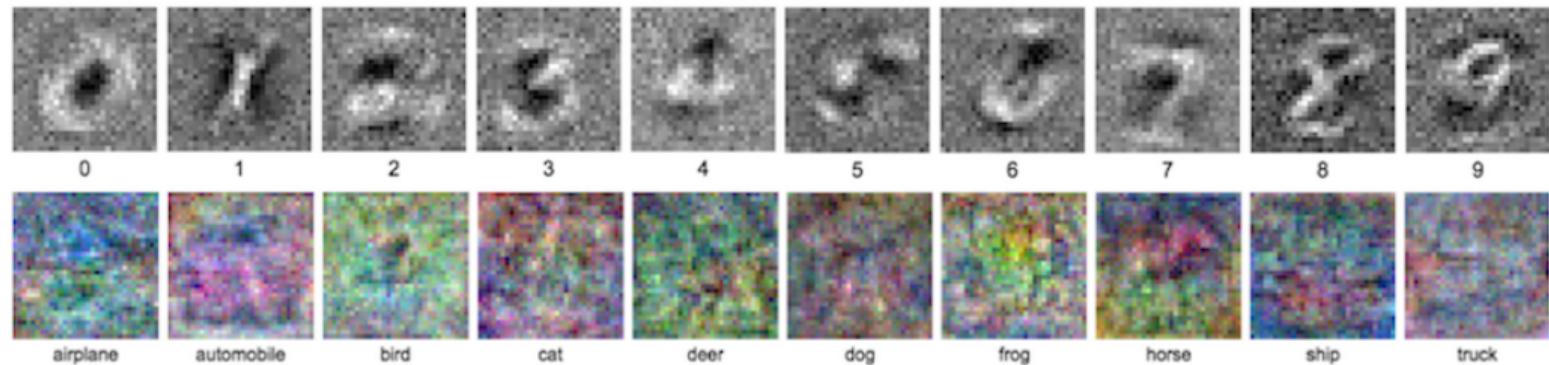
Ульянкин Филипп

13 ноября 2018 г.

Куча разных прикольных архитектур

Немного моих долгов

# Что внутри нейронов



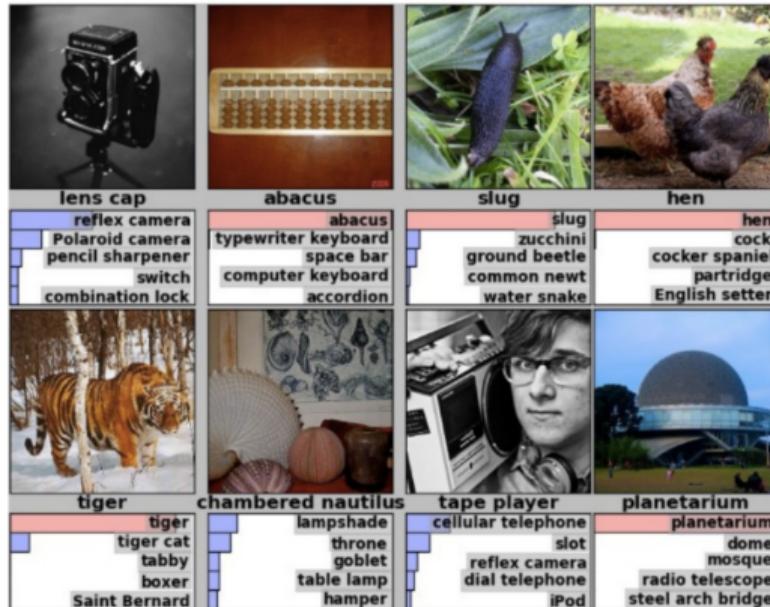
# Сгенерированные стулья



Первая и последняя колонка с реальными стульями, остальное сгенерировано нейросеткой

# Сказ про то как люди ImageNet рвали

# ImageNet



- около 10 миллионов размеченных изображений из интернета



- бывают спорные изображения: тут вишня, если распознать как далматинец, будет неправильно

# ImageNet

- выборка очень неоднородная, на ней проводятся соревнования по распознанию образов
- обычно изображение требуется отнести к одному из 1000 классов, можно давать несколько ответов
- если один из пяти вариантов оказался верным, то классификация считается верной
- до 2012 года лучшие алгоритмы дают ошибку в 25%
- в 2012 году на арену выходят глубокие нейронные сети

# AlexNet



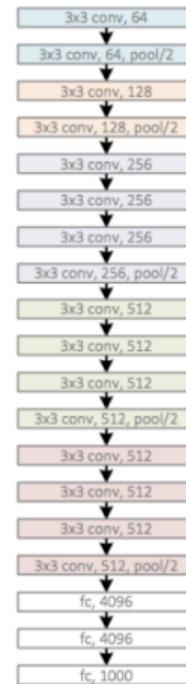
- 5 свёрточных слоёв, два полносвязных слоя, выходной слой  $\Rightarrow 16.5\%$  ошибок

# Ансамбль сетей

- Давайте обучим много сеток
- Каждая сетка будет классифицировать
- Итоговый результат будем выбирать методом большинства
- Ошибка упала до 11.7%

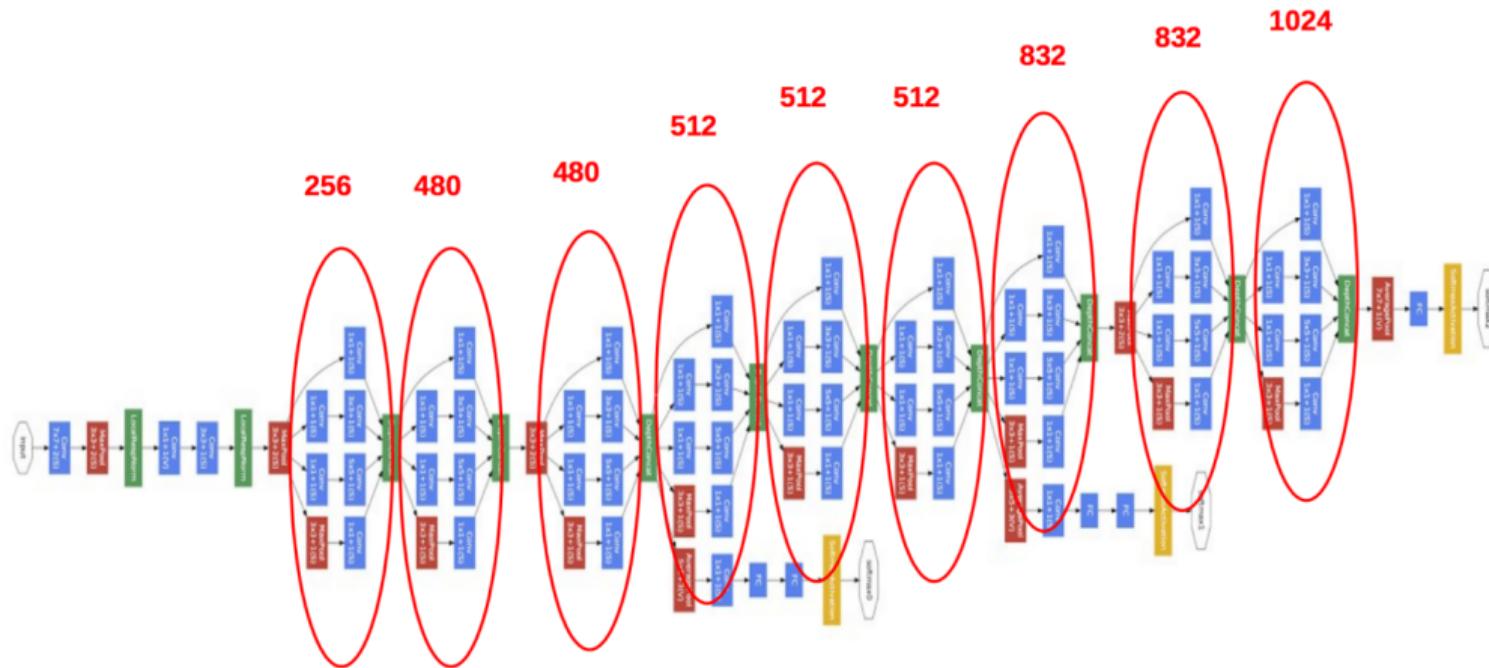
# VGG

VGG, 19 layers  
(ILSVRC 2014)

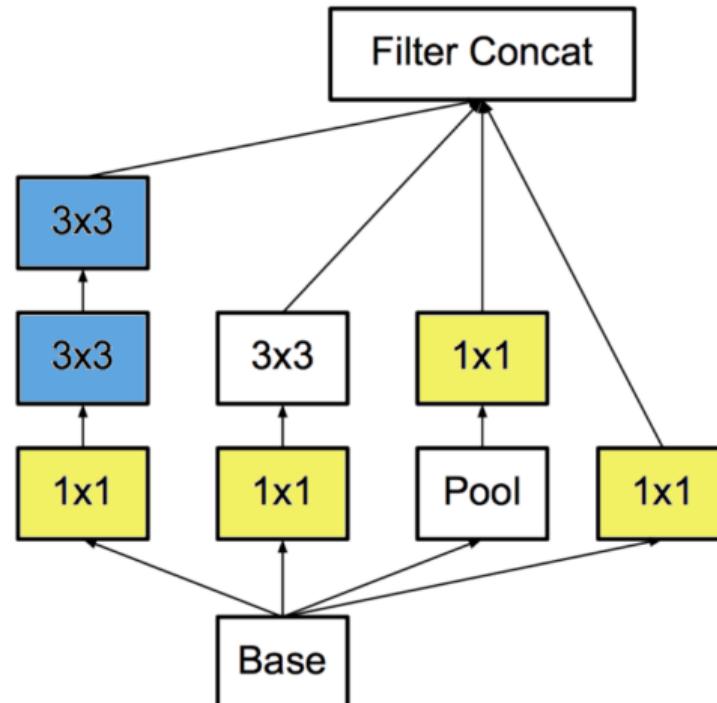


- 19 слоёв, большая часть свёрточные, все свёртки  $3 \times 3 \Rightarrow$  ошибка 7.3%.

# GoogleNet aka Inception V1



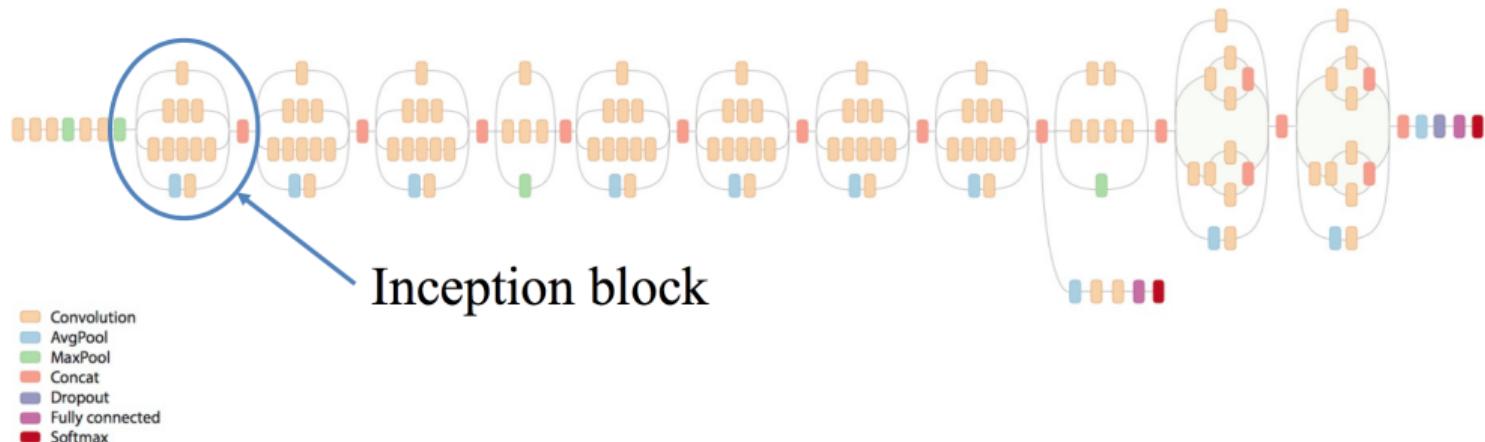
# GoogleNet aka Inception V1



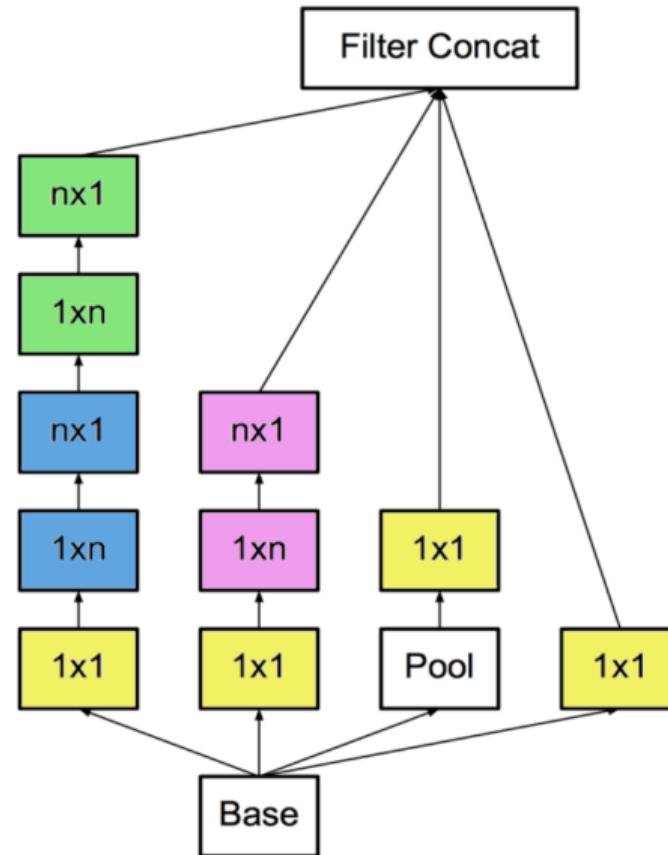
# GoogleNet aka Inception V1

- Сетка собирается из компонент
- На каждом слое используется ни одна свёртка, а несколько разных, что помогает реагировать на сигналы разного масштаба и улучшает работу
- Несколько дополнительных классификаторов на разных уровнях
- идея в том, что такие классификаторы позволяют «протолкнуть» градиенты к ранним слоям и тем самым уменьшить эффект затухания градиента
- Итоговое качество 6.7% ошибок

# Inception V3



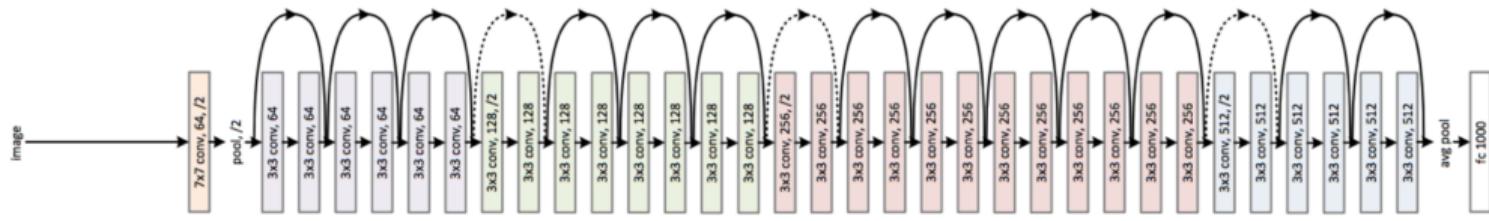
# Inception V3



## Inception V3

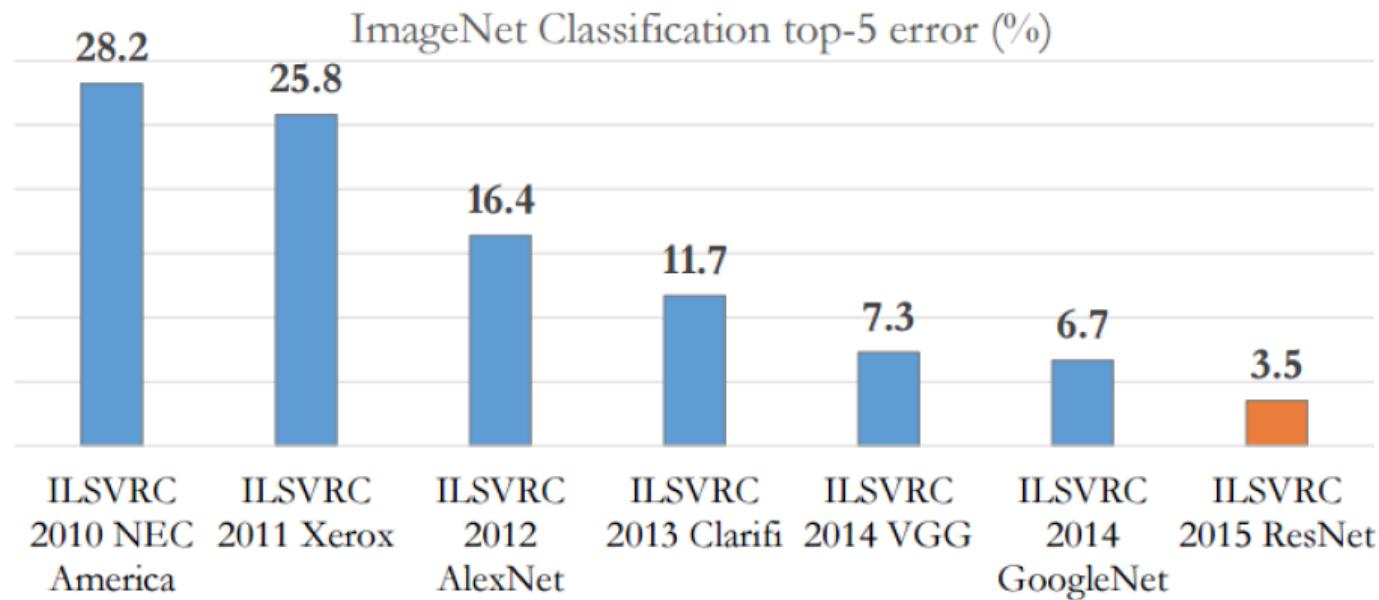
- Сетка собирается из компонент
- На каждом слое используется ни одна свёртка, а несколько разных, что помогает реагировать на сигналы разного масштаба и улучшает работу
- Несколько дополнительных классификаторов на разных уровнях
- идея в том, что такие классификаторы позволяют «протолкнуть» градиенты к ранним слоям и тем самым уменьшить эффект затухания градиента (позже от идеи отказались)
- Итоговое качество 5.6% ошибок

# ResNet (Microsoft)



- 152 слоя, ошибка составила 3.75%
- ключевым элементом архитектуры является связь, которая пропускает несколько слоёв, передавая результат предыдущего слоя
- такое изменение позволило полностью отказаться от таких техник регуляризации, как DropOut

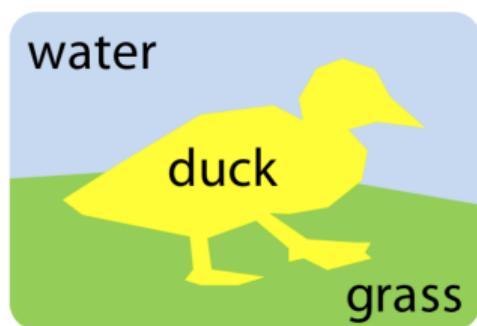
# Итоги



# Сегментация и локализация изображений

# Сегментация и локализация

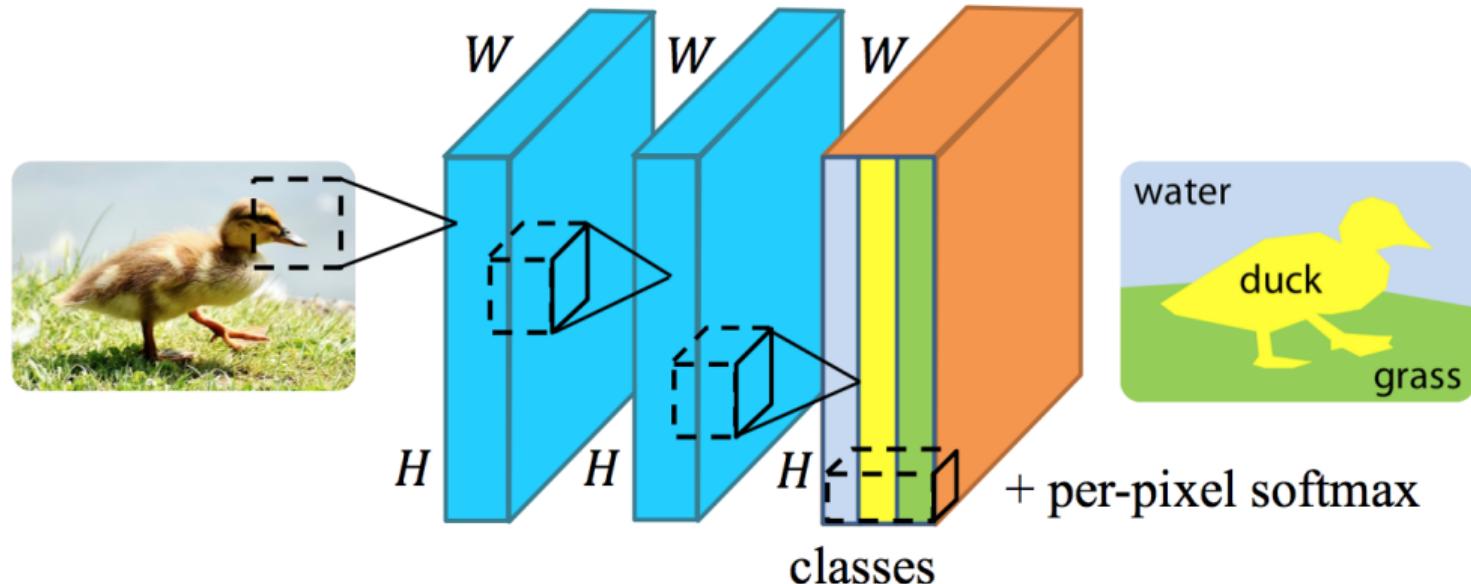
Semantic segmentation:



Object classification  
+ localization:

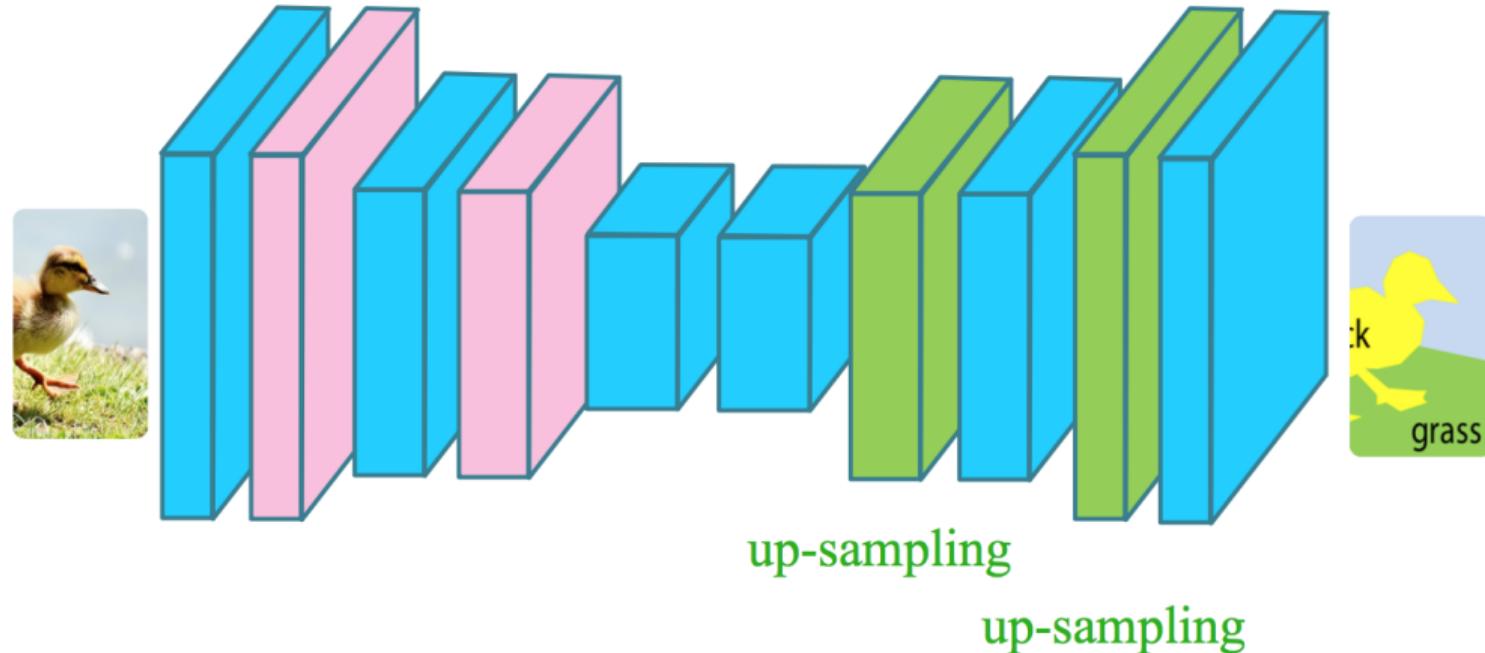


# Сегментация



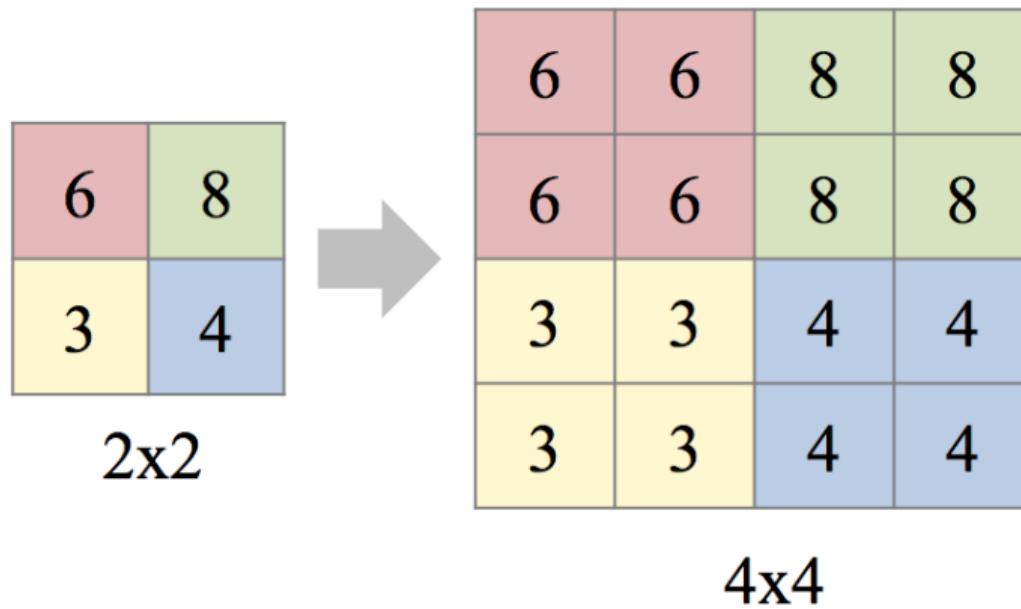
- Нам нужно научиться классифицировать каждый пиксель
- Первая идея: куча свёрток и попиксельный softmax

# Сегментация

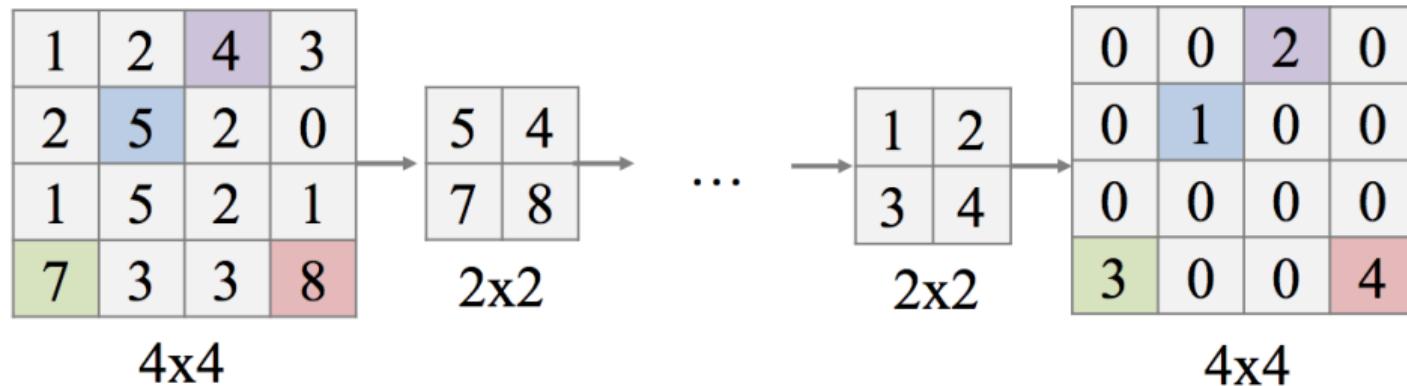


- Если захотим добавить пулинг, придётся делать анпулинг
- Подобную технику мы уже видели в автокодировщиках

# Nearest neighbor unpooling

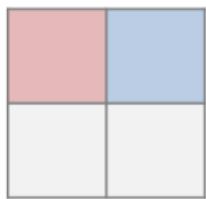


# Max unpooling

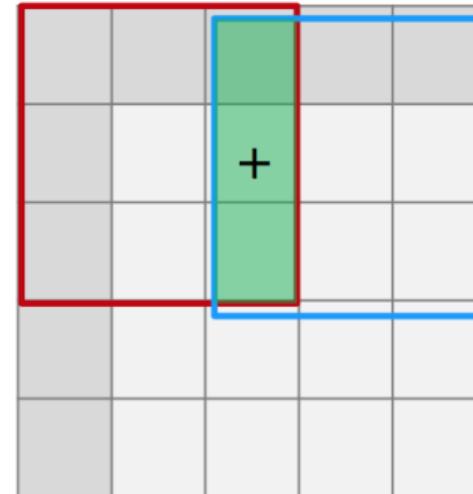


# Learnable unpooling

Input: 2x2



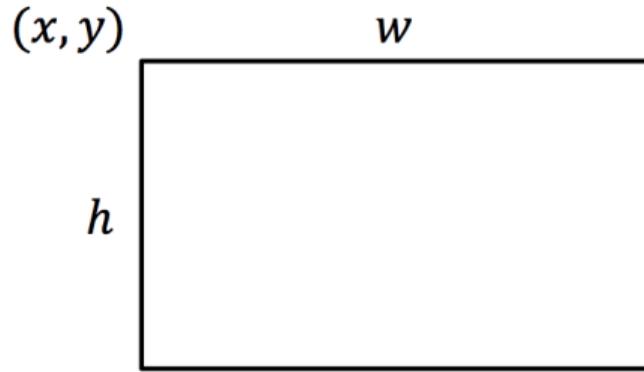
Input gives  
weight for  
filter



Stride: 2

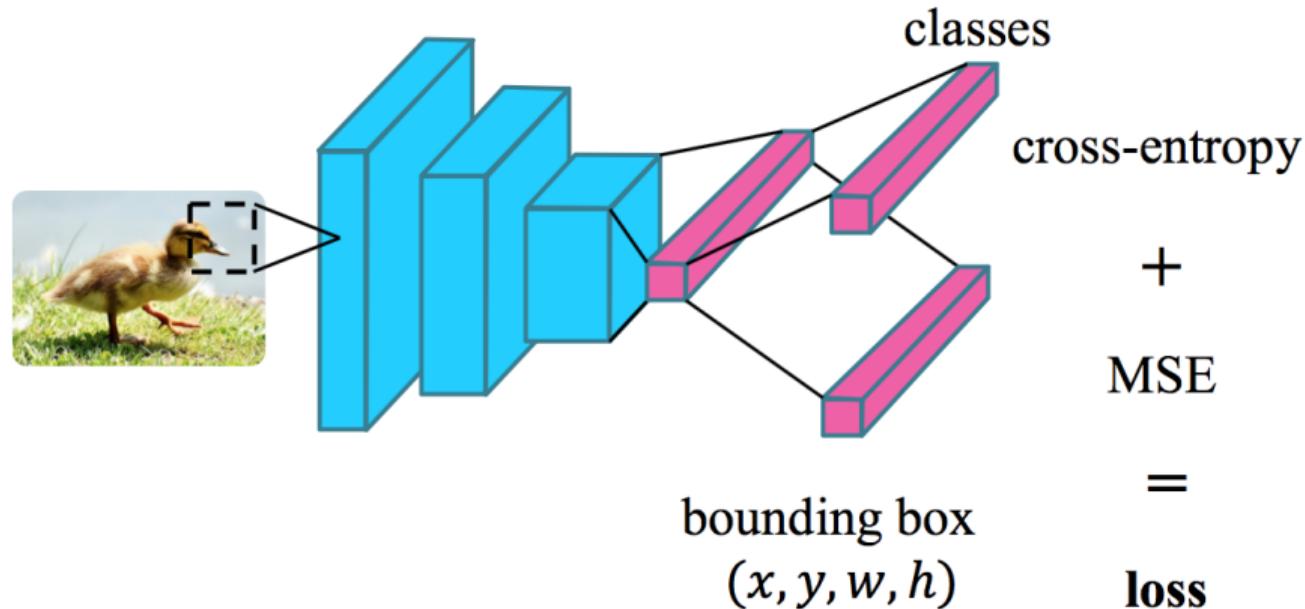
Output: 4x4

# Локализация

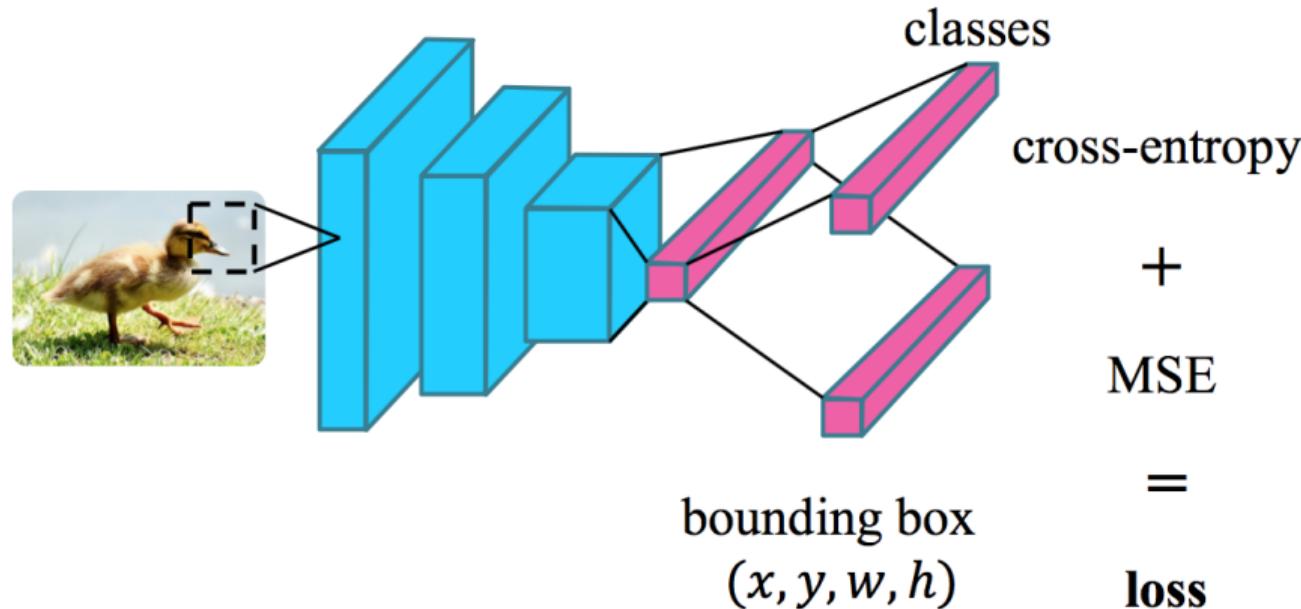


- для локализации объекта нужно нащупать рамочку, в котором он находится
- рамочка описывается параметрами  $(x, y, w, h)$

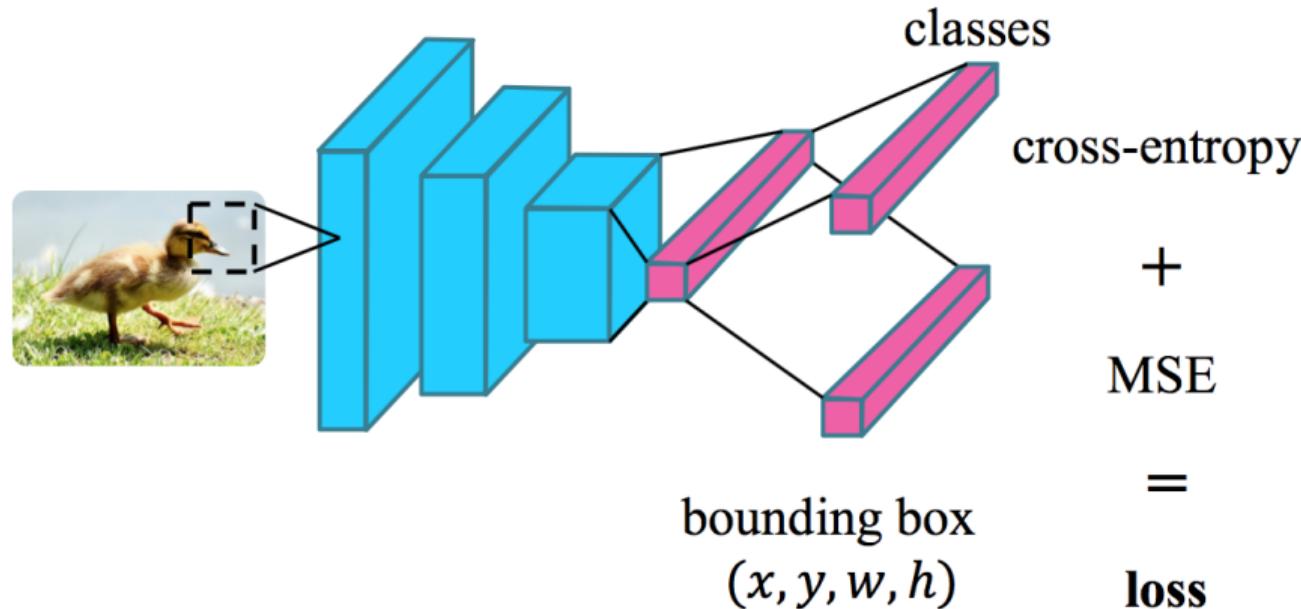
# Локализация



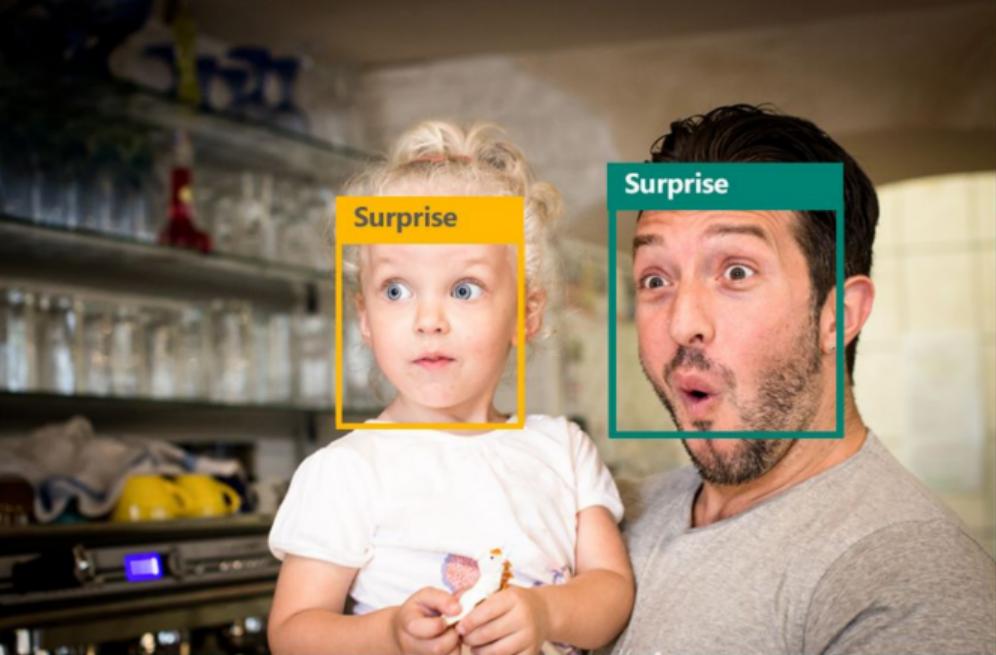
# Примеры



# Примеры



# Примеры



The image shows a young girl with blonde hair and a man with dark hair and a beard, both looking surprised. The girl's face is highlighted with a yellow bounding box, and the word "Surprise" is written above it. The man's face is highlighted with a teal bounding box, and the word "Surprise" is written above it.

Neutral:  

Happiness:  

Surprise:  

Sadness:  

Anger:  

Disgust:  

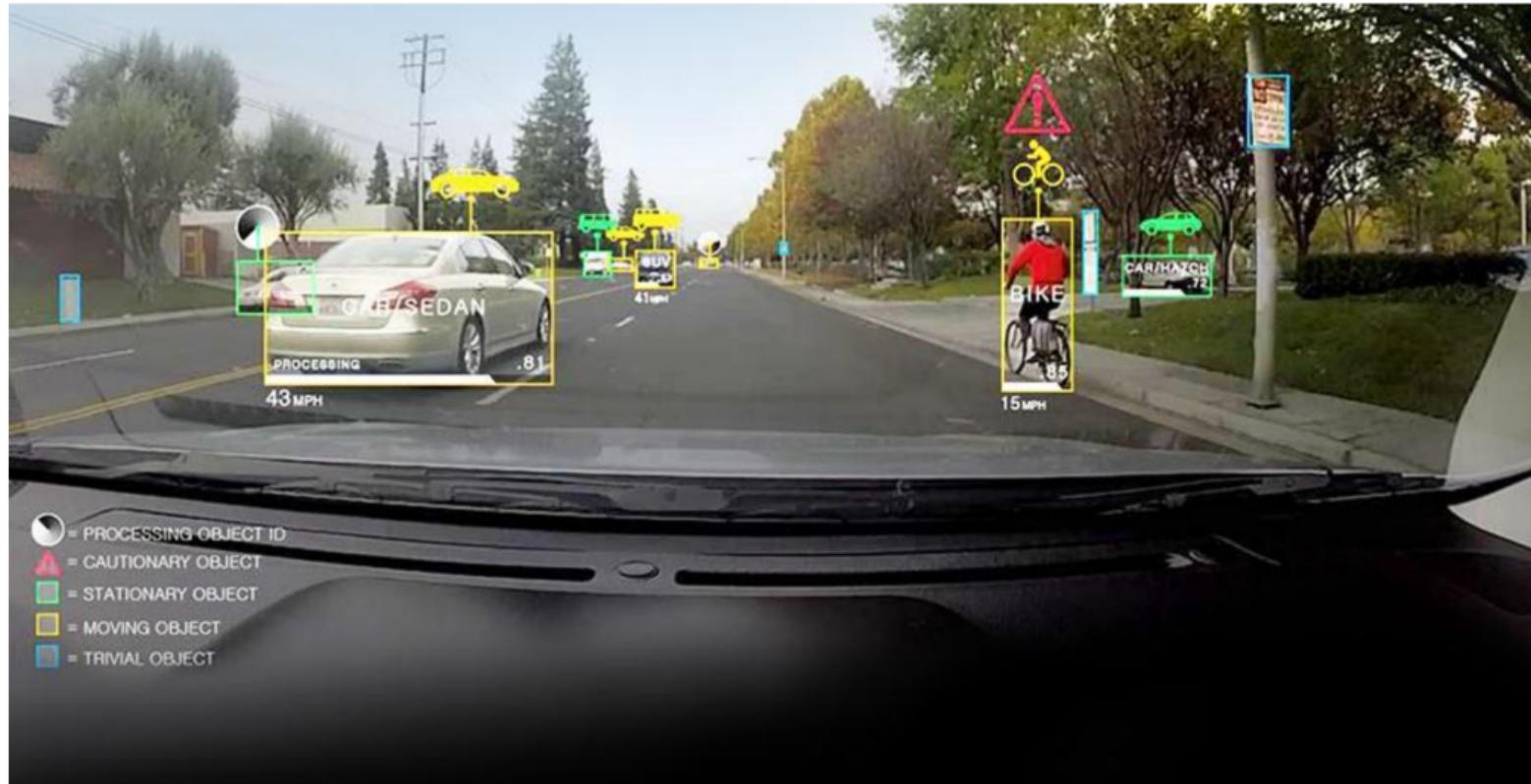
Fear:  

Contempt:  

 Microsoft

Get started for free at [projectoxford.ai](http://projectoxford.ai)

# Примеры



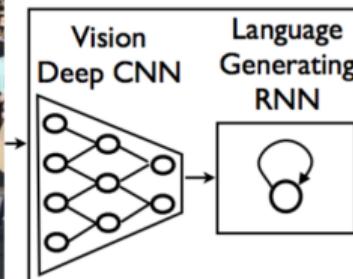
# Примеры



<https://www.youtube.com/watch?v=ZJMtDRbqH40>

Текст + картинка

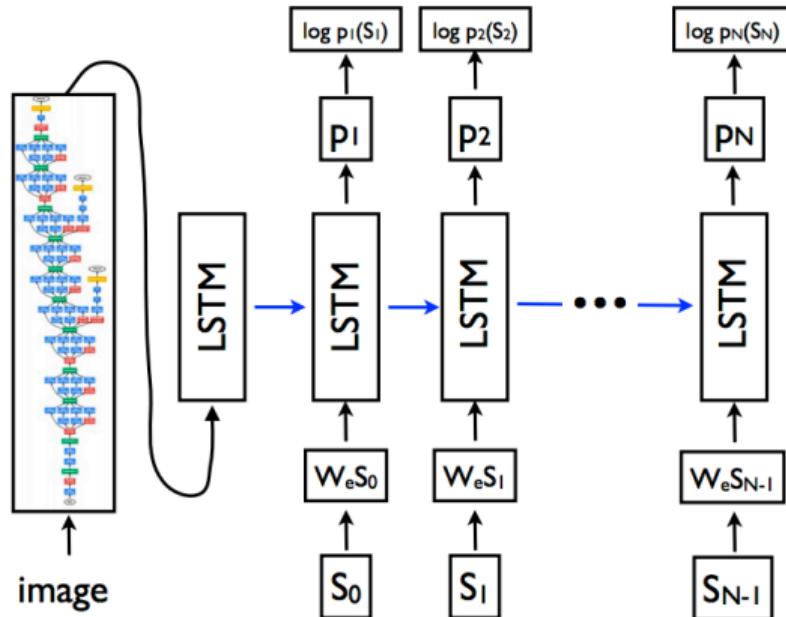
# Генерация подписи по картинке



**A group of people  
shopping at an  
outdoor market.**

**There are many  
vegetables at the  
fruit stand.**

# Генерация подписи по картинке



<https://arxiv.org/abs/1411.4555>

# Генерация подписи по картинке

A person riding a motorcycle on a dirt road.



Two dogs play in the grass.



A skateboarder does a trick on a ramp.



A dog is jumping to catch a frisbee.



A group of young people playing a game of frisbee.



Two hockey players are fighting over the puck.



A little girl in a pink hat is blowing bubbles.



A refrigerator filled with lots of food and drinks.



A herd of elephants walking across a dry grass field.



A close up of a cat laying on a couch.



A red motorcycle parked on the side of the road.



A yellow school bus parked in a parking lot.



Describes without errors

Describes with minor errors

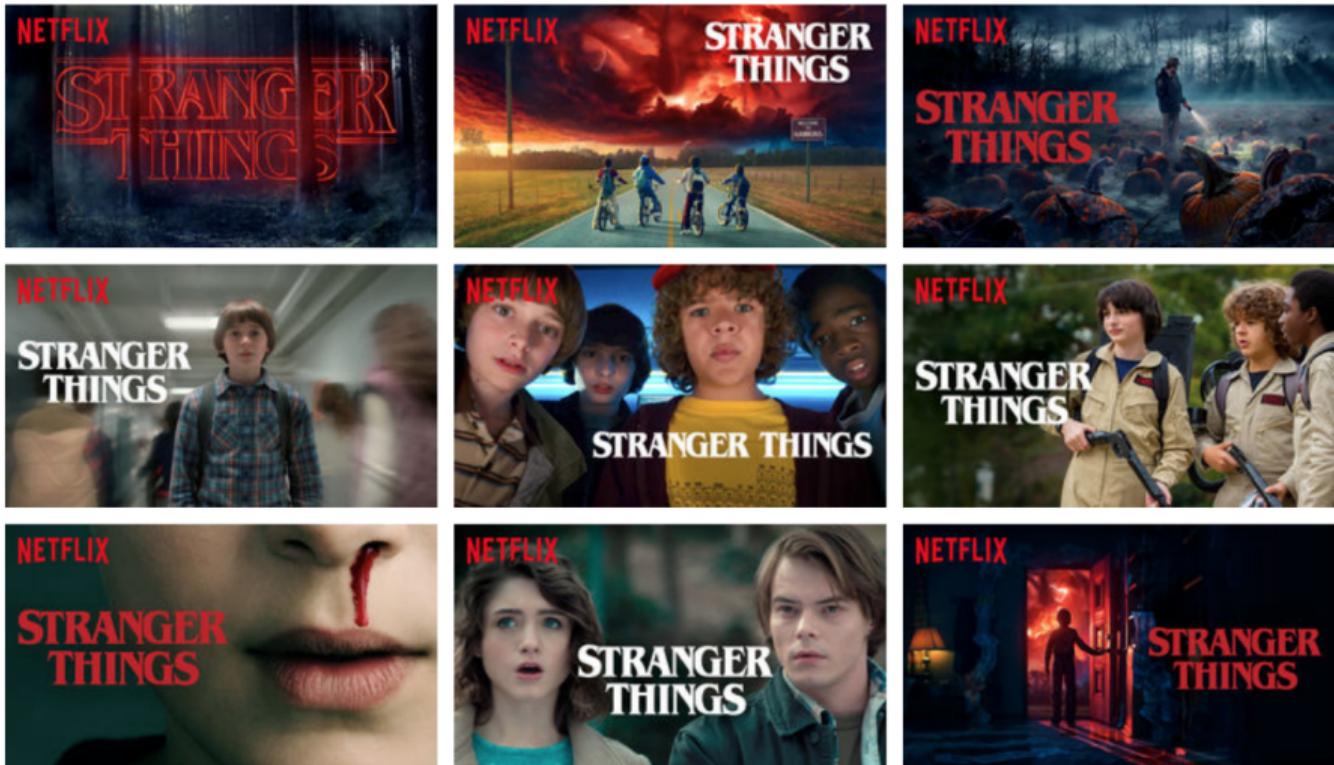
Somewhat related to the image

Unrelated to the image

Figure 5. A selection of evaluation results, grouped by human rating.

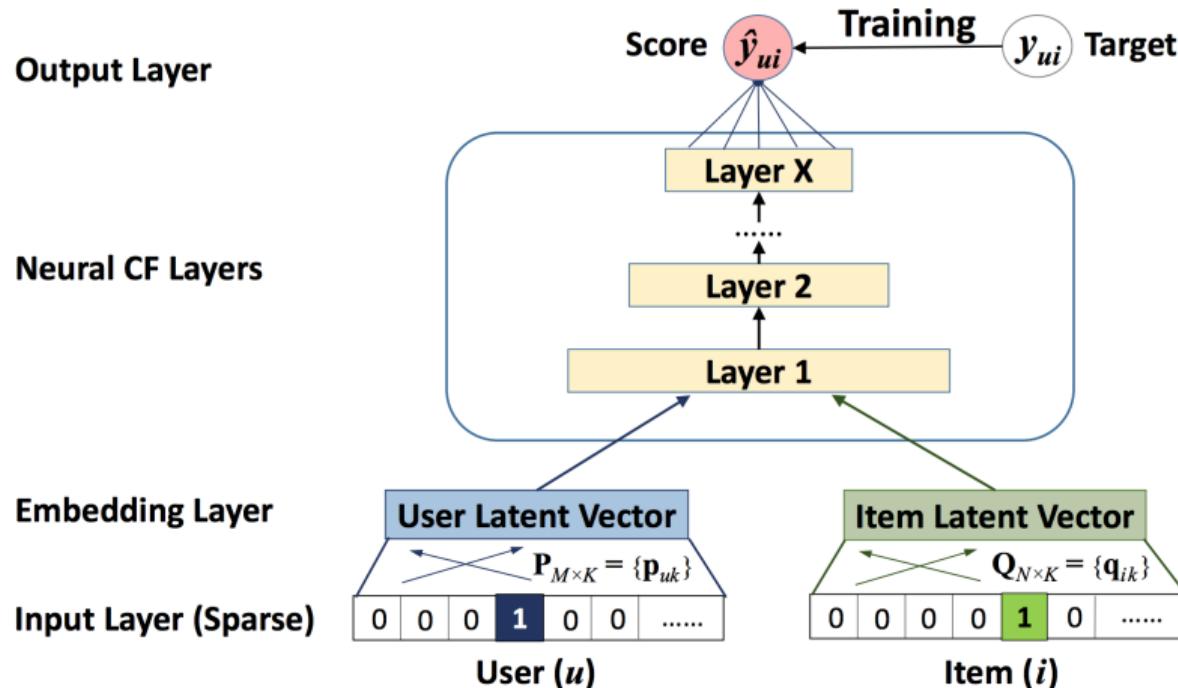
# Рекомендательные системы и w2v

# Рекомендации - это важно



<https://medium.com/netflix-techblog/artwork-personalization-c589f074ad76>

# Neural Collaborative Filtering



# Нейробайесовские методы

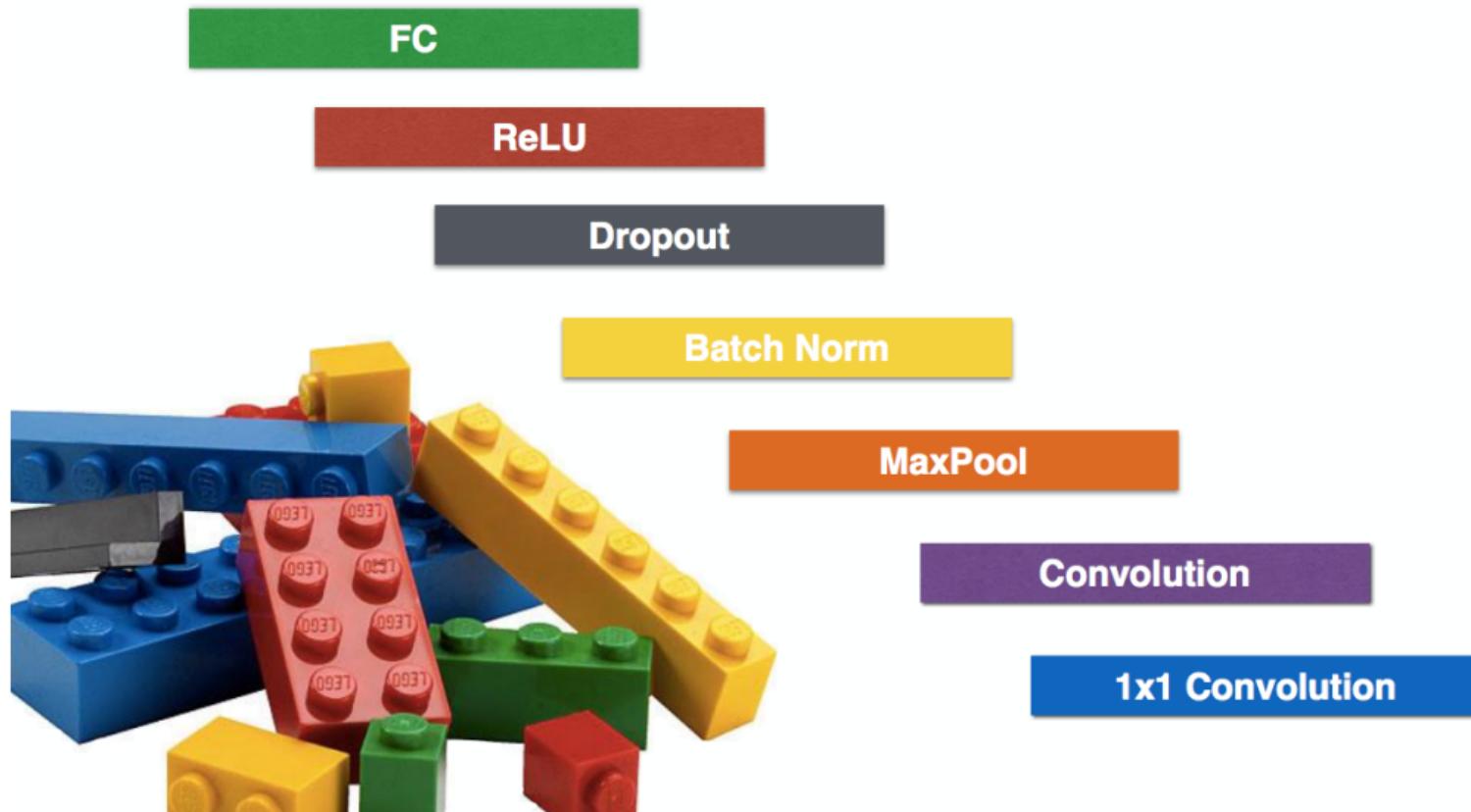
# Байесовский w2v

- ватерло - лондон - станция - поезд
- ватерло - наполеон - аустерлиц - битва
- ватерло - абба - мама-мия

## Байесовский w2v

- S. Bartunov, D. Kondrashkin, A. Osokin, D. Vetrov. Breaking Sticks and Ambiguities with Adaptive Skip-gram. In AISTATS 2016  
<http://arxiv.org/abs/1502.07257>
- Код и документация: <https://github.com/sbos/AdaGram.jl>
- Предобученная модель: <https://yadi.sk/d/W4FtSjA5o3jUL>

# Не забывайте, что нейросети - Lego



## Не забывайте искать

- Перед тем как решать задачу, убедитесь, что уже нет решения
- Один из зоопарков моделей:  
<https://github.com/BVLC/caffe/wiki/Model-Zoo>