



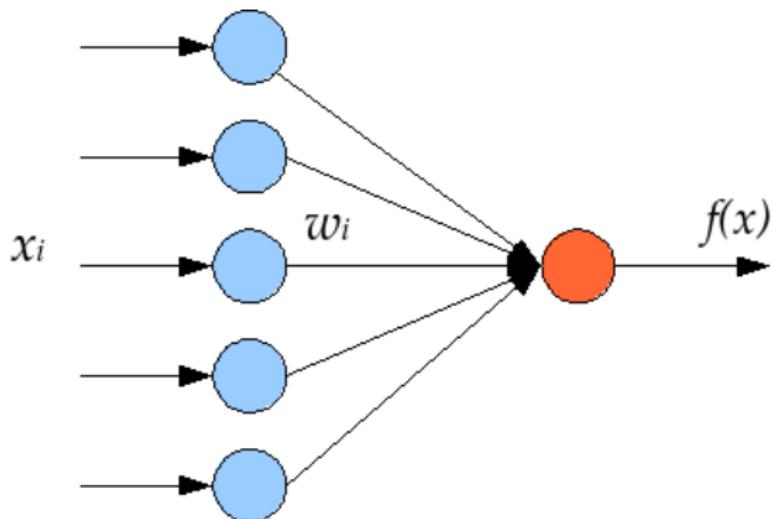
# ТЕХНОСФЕРА

## Лекция 4 Сверточные нейронные сети

Полыковский Даниил

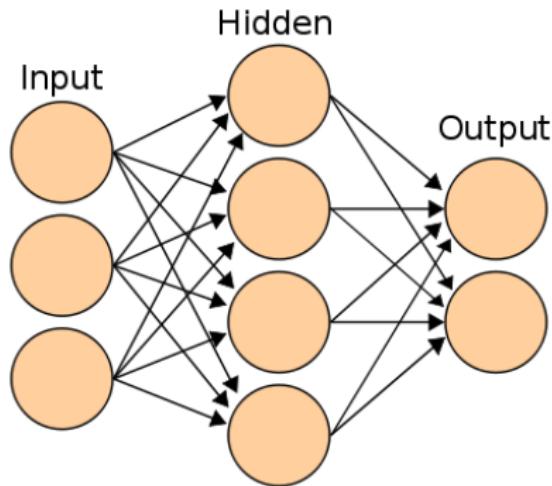
2 октября 2017 г.

# Перцептрон



- ▶ Модели: линейная/логистическая регрессия
- ▶ Может моделировать: NOT, AND, OR
- ▶ Не может моделировать: XOR

## Сети с одним скрытым слоем



Теорема (универсальный аппроксиматор)<sup>1</sup>

Любую непрерывную на компакте функцию можно равномерно приблизить нейронной сетью с одним скрытым слоем.

---

<sup>1</sup>Отличная визуализация:

<http://neuralnetworksanddeeplearning.com/chap4.html>

# Проблемы нейронных сетей

Проблемы полносвязных нейронных сетей:

- ▶ Требуется огромное количество нейронов
- ▶ Серьезное переобучение

Возможное решение — введение новых типов слоев:

- ▶ Сверточные слои (сегодня)
- ▶ Пулинг (сегодня)
- ▶ Dropout (лекция 6)
- ▶ Нормализация (лекция 6)
- ▶ ...

# Сверточные нейронные сети

# ImageNet



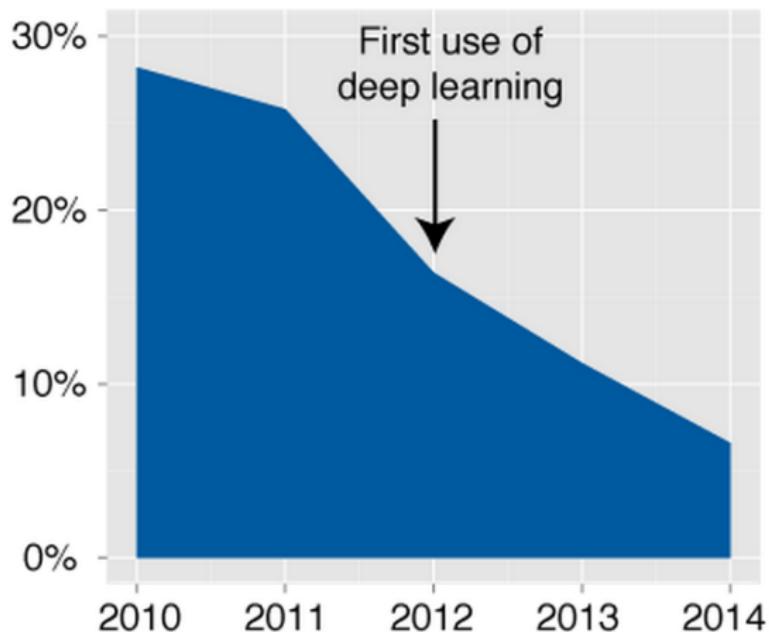
- ▶ 1000 классов
- ▶ около 1000 изображений в каждом классе
- ▶ около 1 000 000 изображений всего
- ▶ несколько номинаций: таких как распознавание и детектирование/локализация

# ImageNet

|   |  |  |   |
|---|--|--|---|
|  |           |       |           |
| <b>mite</b><br>black widow<br>cockroach<br>tick<br>starfish                       | <b>container ship</b><br>lifeboat<br>amphibian<br>fireboat<br>drilling platform            | <b>motor scooter</b><br>go-kart<br>moped<br>bumper car<br>golfcart                     | <b>leopard</b><br>jaguar<br>cheetah<br>snow leopard<br>Egyptian cat                         |
|  |           |       |           |
| <b>grille</b><br>convertible<br>grille<br>pickup<br>beach wagon<br>fire engine    | <b>mushroom</b><br>agaric<br>mushroom<br>jelly fungus<br>gill fungus<br>dead-man's-fingers | <b>cherry</b><br>dalmatian<br>grape<br>elderberry<br>ffordshire bullterrier<br>currant | <b>Madagascar cat</b><br>squirrel monkey<br>spider monkey<br>titi<br>indri<br>howler monkey |

Рис.: Примеры прогнозов

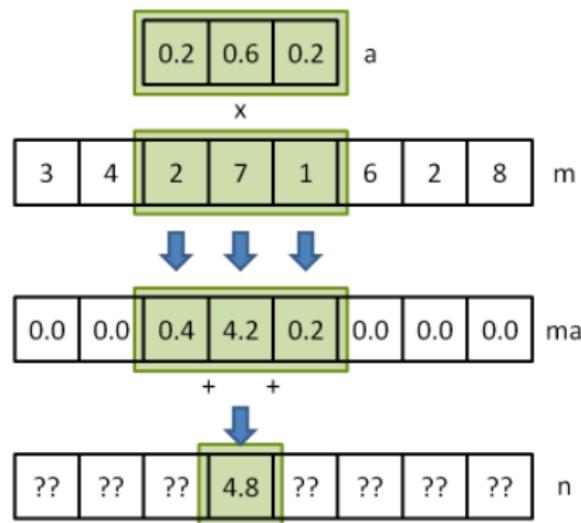
## Objection classification error rate



# Одномерная свертка (convolution)

## Определение

Результатом операции свертки массива  $m$  с ядром  $a$  называется сигнал  $n$ :  $n[k] = \sum_{i=-w}^w m[k+i]a[-i]$ . Обозначение:  $n = m * a$



# Padding

Нулевой отступ

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| 0 | 0 | A | B | C | 0 | 0 |
|---|---|---|---|---|---|---|

Продолжение границы

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| A | A | A | B | C | C | C |
|---|---|---|---|---|---|---|

Зеркальный отступ

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| B | A | A | B | C | C | B |
|---|---|---|---|---|---|---|

|   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
| C | B | A | B | C | C | B | A |
|---|---|---|---|---|---|---|---|

Циклический отступ

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| B | C | A | B | C | A | B |
|---|---|---|---|---|---|---|

# Одномерная свертка (convolution)

Свойства (для сигналов бесконечной длины):

- ▶ Ассоциативность:  $a * (b * c) = (a * b) * c$
- ▶ Линейность:
  - ▶  $(\alpha a) * b = \alpha(a * b)$
  - ▶  $(a + b) * c = a * c + b * c$
- ▶ Коммутативность:  $a * b = b * a$

# Двумерная свертка (чб картинки)

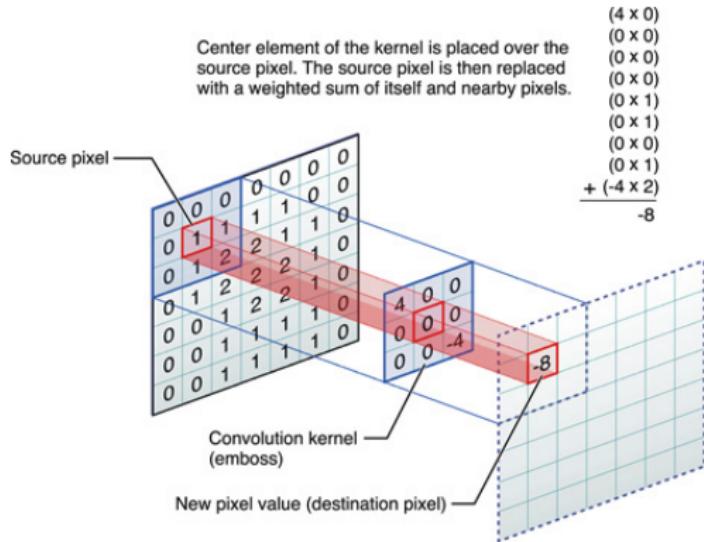


Рис.: 2D convolution<sup>2</sup>

<sup>2</sup><https://developer.apple.com/library/ios/documentation/Performance/Conceptual/vImage/ConvolutionOperations/ConvolutionOperations.html>

# Примеры ядер

- ▶ Тождественное

|   |   |   |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 0 | 0 | 0 |



- ▶ Детектор границ

|   |    |   |
|---|----|---|
| 0 | 1  | 0 |
| 1 | -4 | 1 |
| 0 | 1  | 0 |



- ▶ Увеличение резкости

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



# Свертка в нейронных сетях

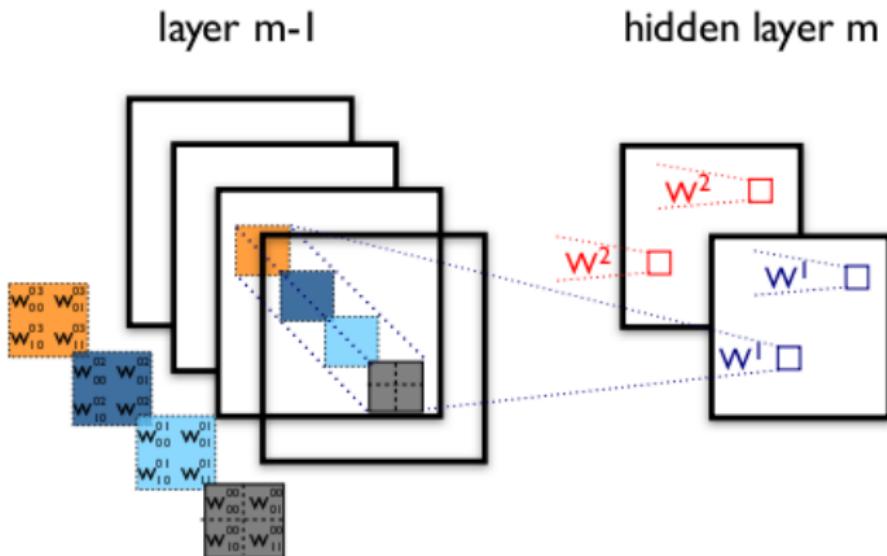
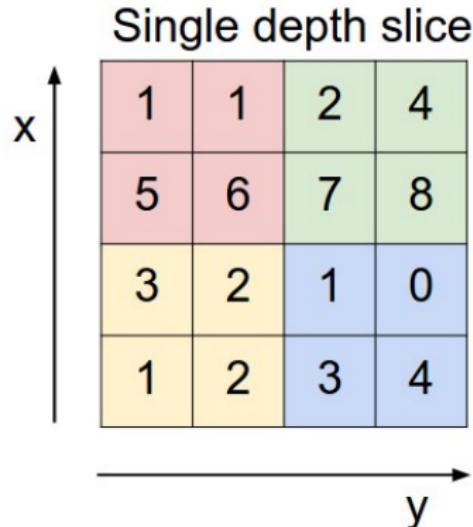


Рис.: Сверточный слой<sup>3</sup>

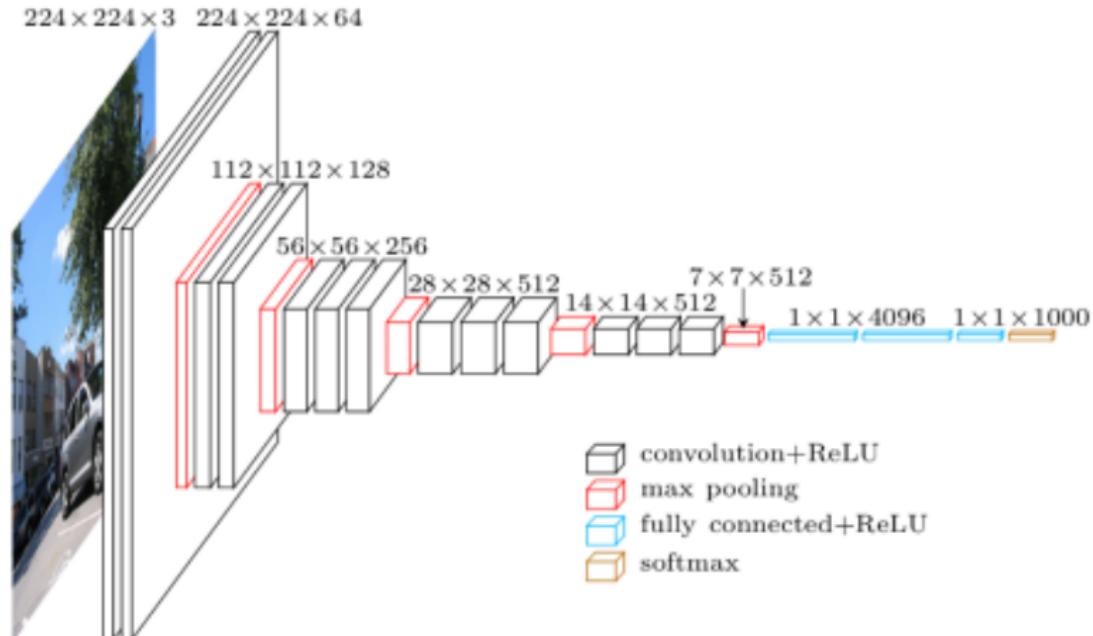
<sup>3</sup><http://deeplearning.net/tutorial/lenet.html>

# Pooling



- ▶ Голосование: побеждают наиболее активные нейроны
- ▶ Вырабатывается инвариантность к небольшим сдвигам
- ▶ Увеличение рецептивной области
- ▶ Уменьшение вычислительных затрат
- ▶ Кроме max-пулинга: mean, weighted, root-mean-square, ...

# Пример: VGG-16



# Интерпретация обученных моделей

# Извлечение признаков

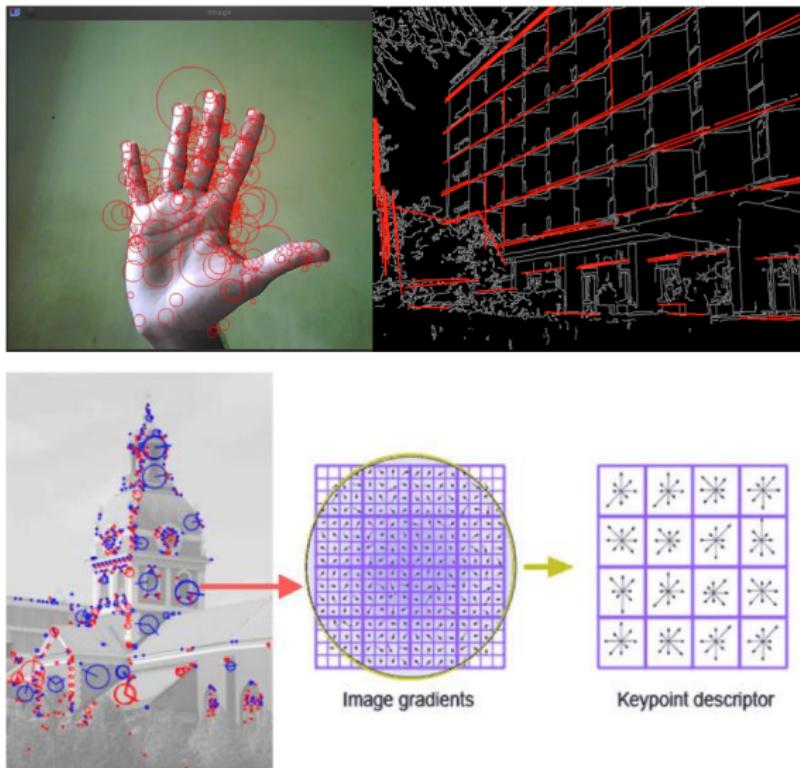


Рис.: Классический подход к извлечению признаков

# Извлечение признаков, история

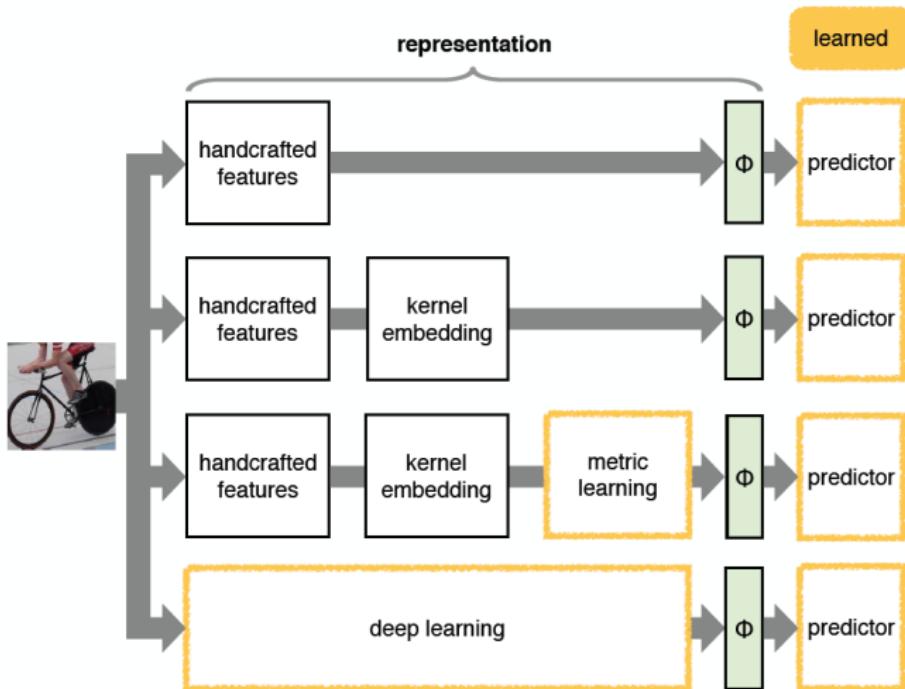
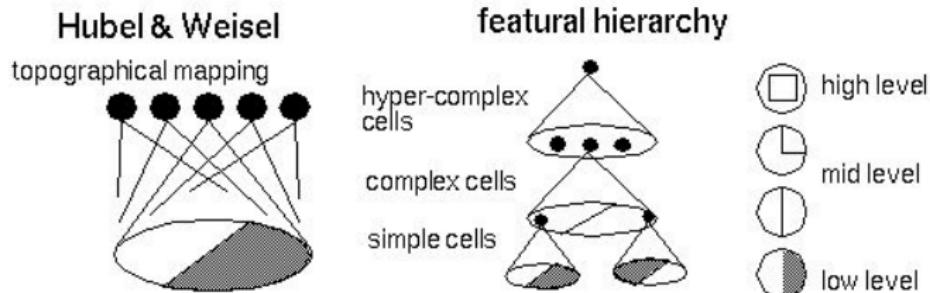


Рис.: Глубинное обучение<sup>4</sup>

<sup>4</sup>Learning visual representations (Andrea Vedaldi)

# Модель Хьюбеля-Визеля



Показано, что мозг обрабатывает визуальную информацию иерархически: сначала находятся границы, углы, а на более глубоких слоях — сложные объекты.

# Deconvolution сети

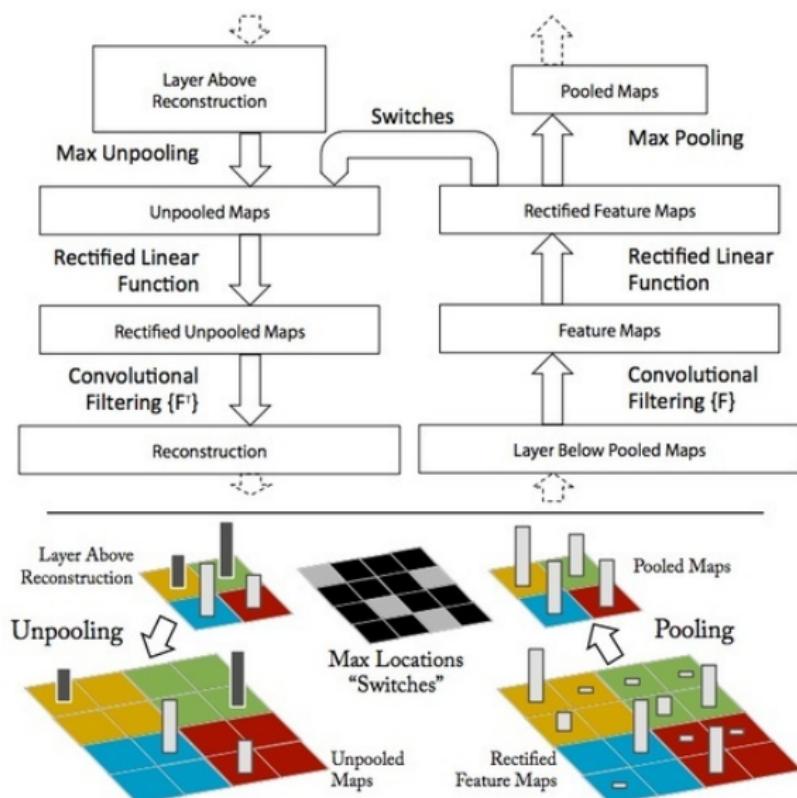


Рис.: Схема deconvolution сети

# Deconvolution сети

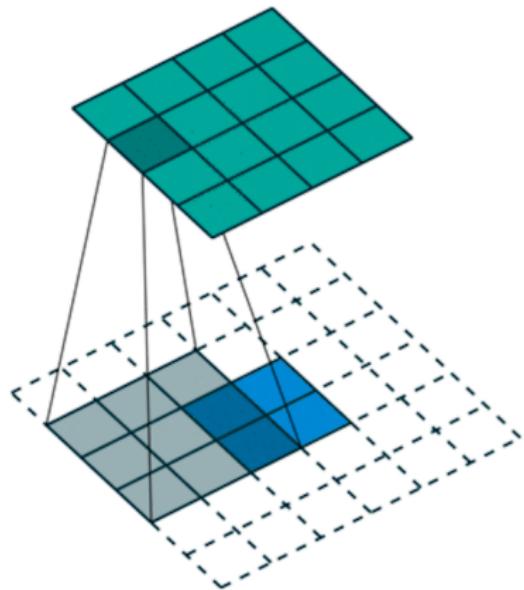


Рис.: Convolution transposed<sup>5</sup>

---

<sup>5</sup>[https://github.com/vdumoulin/conv\\_arithmetic](https://github.com/vdumoulin/conv_arithmetic)

# Deconvolution сети

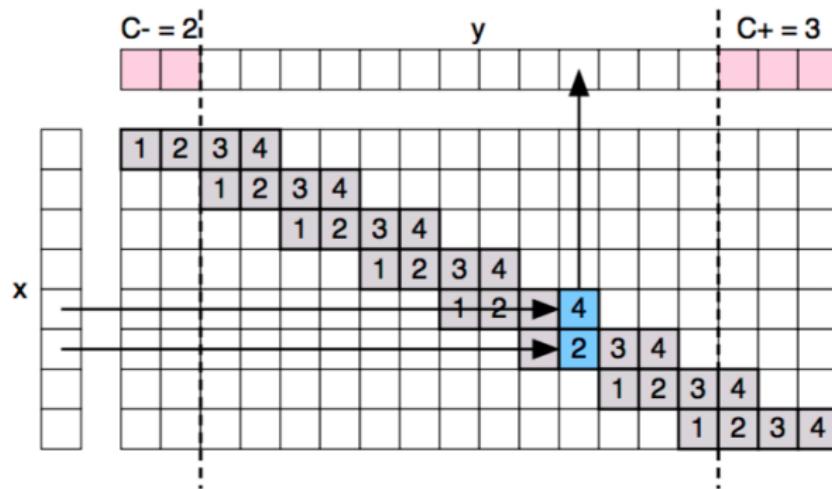


Рис.: Convolution transposed<sup>6</sup>

<sup>6</sup><http://www.vlfeat.org/matconvnet/matconvnet-manual.pdf>

# Выучиваемые признаки

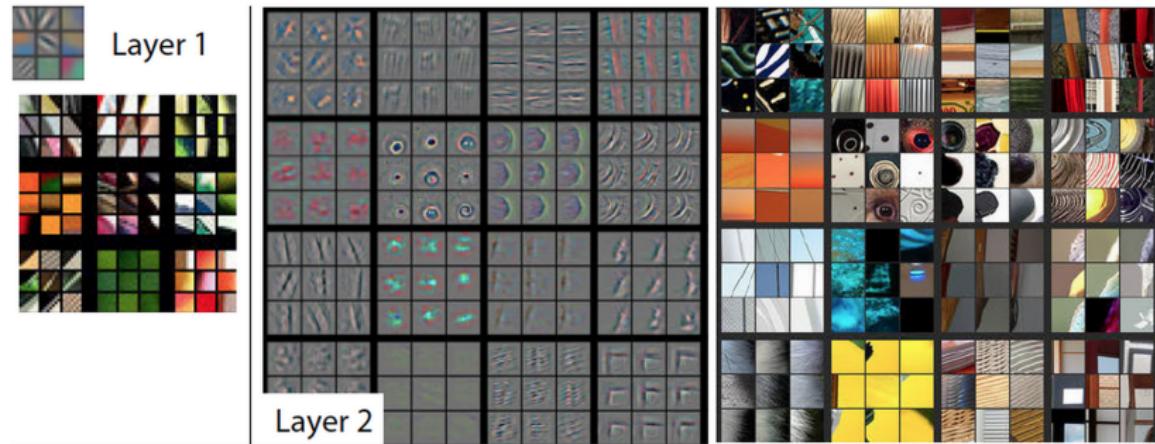


Рис.: Visualizing and Understanding Convolutional Networks<sup>7</sup>

<sup>7</sup>Matthew D. Zeiler and Rob Fergus

# Выучиваемые признаки

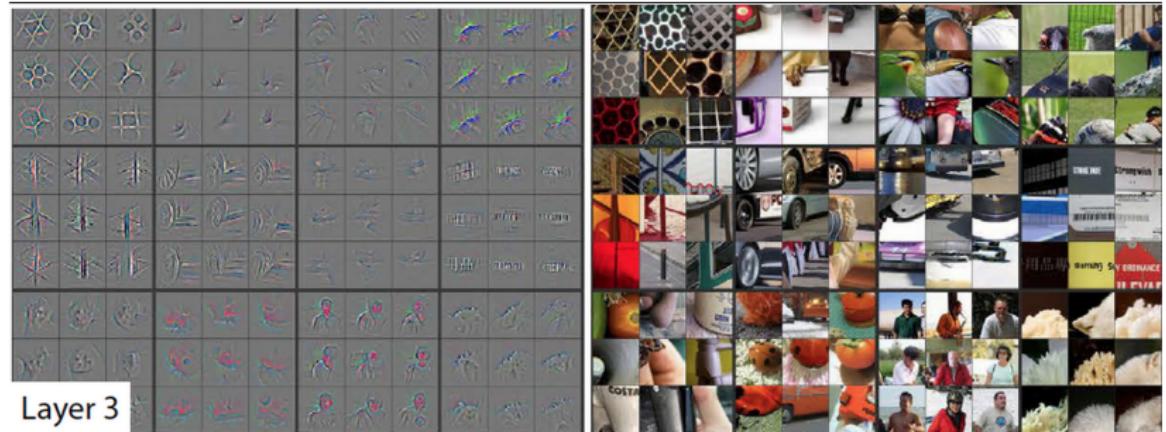


Рис.: Visualizing and Understanding Convolutional Networks<sup>8</sup>

<sup>8</sup>Matthew D. Zeiler and Rob Fergus

# Выучиваемые признаки

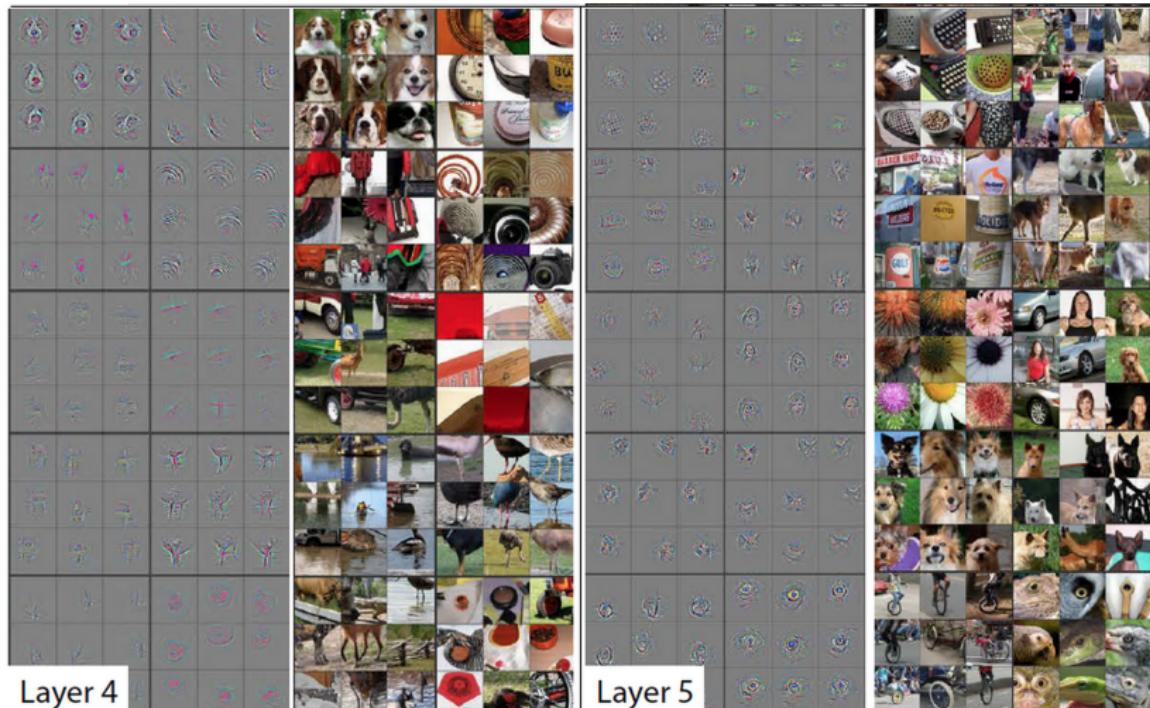


Рис.: Visualizing and Understanding Convolutional Networks<sup>9</sup>

<sup>9</sup>Matthew D. Zeiler and Rob Fergus

# Transfer learning

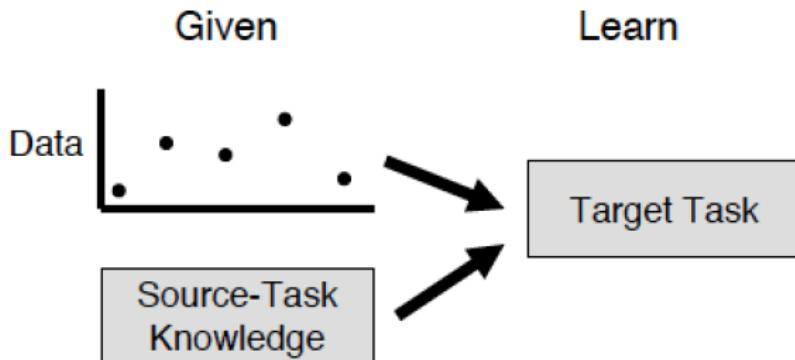


Рис.: Модель решения задачи в рамках парадигмы трансфера знаний

# Transfer learning

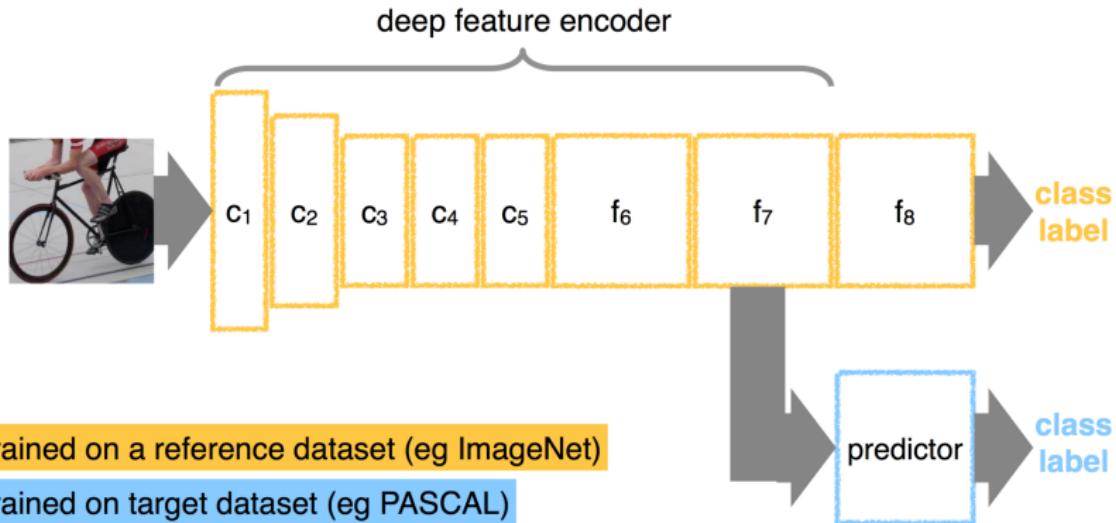


Рис.: Трансфер между двумя глубинными сетями<sup>10</sup>

<sup>10</sup>Learning visual representations (Andrea Vedaldi)

# Transfer learning<sup>11</sup>

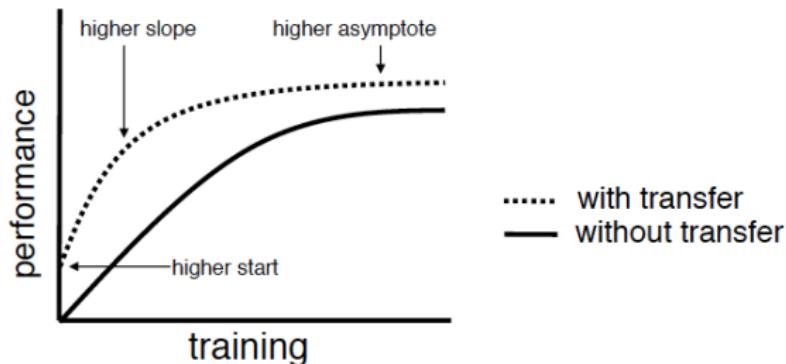


Рис.: Цели трансфера знаний

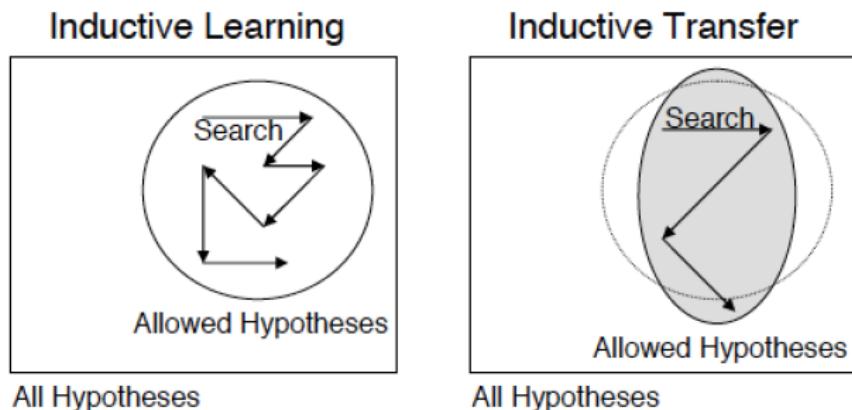
- ▶ higher start — хорошее начальное приближение из-за априорной информации о распределении весов
- ▶ higher slope — ускорение сходимости алгоритма обучения
- ▶ higher asymptote — улучшение верхней достижимой границы качества

---

<sup>11</sup>ftp:

//ftp.cs.wisc.edu/machine-learning/shavlik-group/torrey.handbook09.pdf

# Transfer learning



**Рис.:** Трансфер знаний можно также рассматривать как некоторую регуляризацию, которая ограничивает пространство поиска до определенного набора допустимых и хороших гипотез

# CNN для распознавания речи<sup>12</sup>

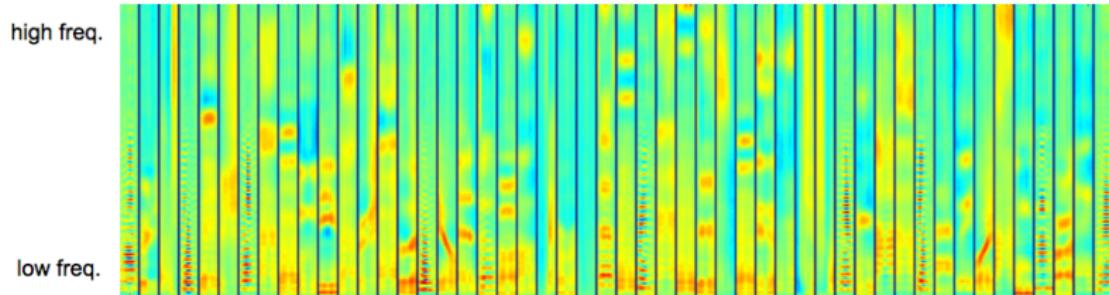


Рис.: Спектрограмма голосового сигнала

---

<sup>12</sup><http://ai.stanford.edu/~ang/papers/nips09-AudioConvolutionalDBN.pdf>

# CNN для текстов<sup>1314</sup>

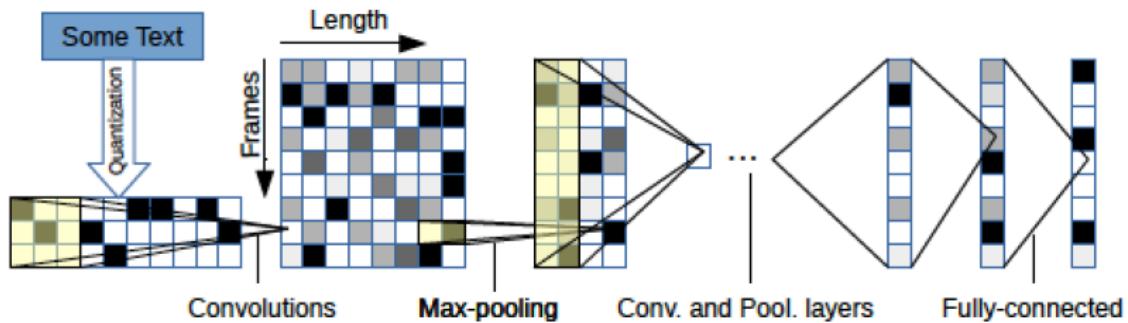


Рис.: Обработка изображения представляющего текст

<sup>13</sup><http://arxiv.org/pdf/1502.01710v1.pdf>

<sup>14</sup>[http://nlp.csail.mit.edu/papers/Kalchbrenner\\_DCNN\\_ACL14](http://nlp.csail.mit.edu/papers/Kalchbrenner_DCNN_ACL14)

## Вопросы

