



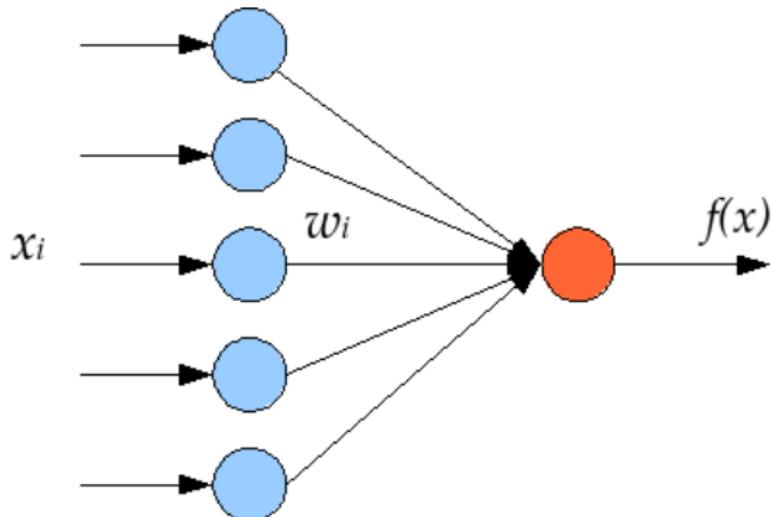
ТЕХНОСФЕРА

Лекция 4 Сверточные нейронные сети

Байгушев Данила

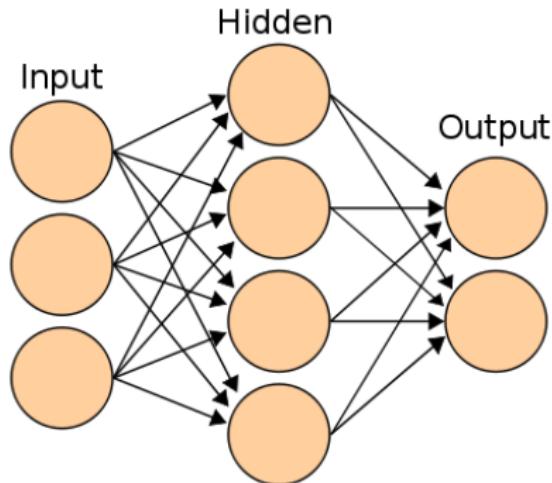
15 сентября 2019 г.

Перцептрон



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

Сети с одним скрытым слоем



Теорема (универсальный аппроксиматор)¹

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

¹Отличная визуализация:

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

Проблемы нейронных сетей

Проблемы полносвязных нейронных сетей:

- ▶ Требуется огромное количество нейронов
- ▶ Серьезное переобучение

Проблемы нейронных сетей

Проблемы полносвязных нейронных сетей:

- ▶ Требуется огромное количество нейронов
- ▶ Серьезное переобучение

Возможное решение — введение новых типов слоев:

- ▶ Сверточные слои (сегодня)
- ▶ Пуллинг (сегодня)
- ▶ Dropout (лекция 5)
- ▶ Нормализация (лекция 5)
- ▶ ...

Сверточные нейронные сети

ImageNet



- ▶ 1000 классов
- ▶ около 1000 изображений в каждом классе
- ▶ около 1 000 000 изображений всего
- ▶ несколько номинаций, в том числе распознавание и детектирование/локализация

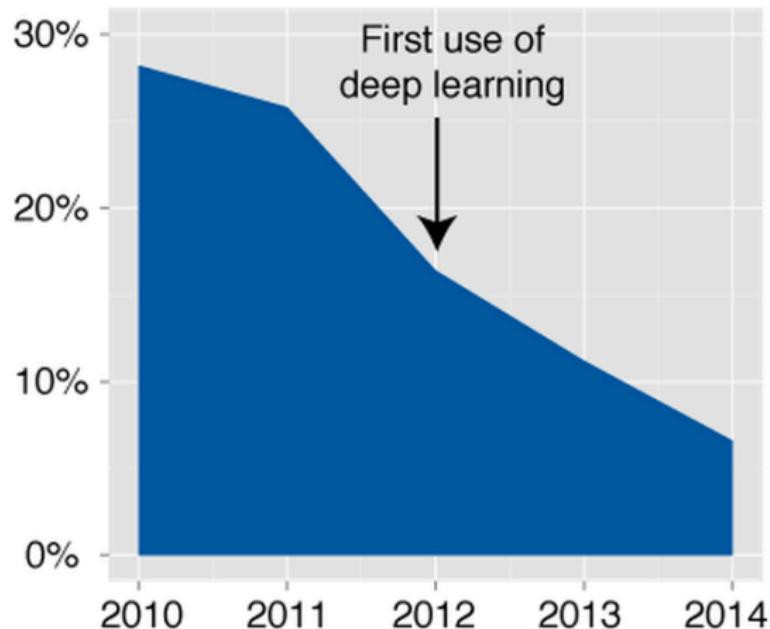
ImageNet

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

Figure: Примеры прогнозов

ImageNet

Objection classification error rate

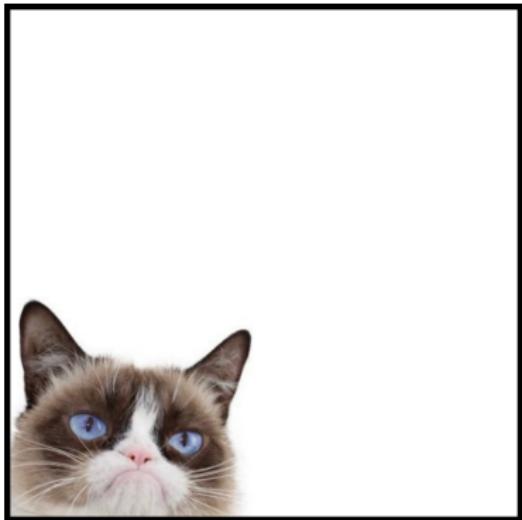


Внутренние инварианты



(a) Кот

Внутренние инварианты

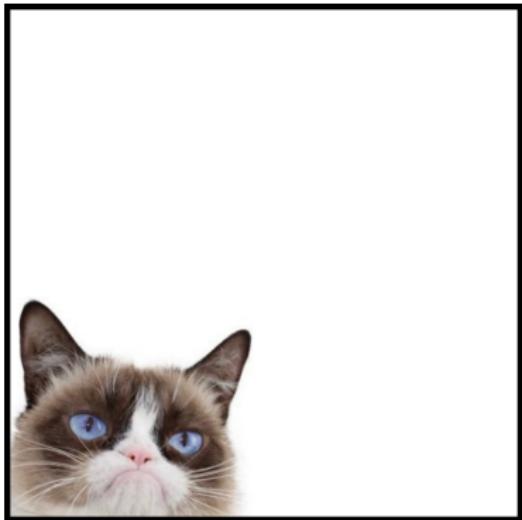


(a) Кот

Внутренние инварианты



(a) Кот

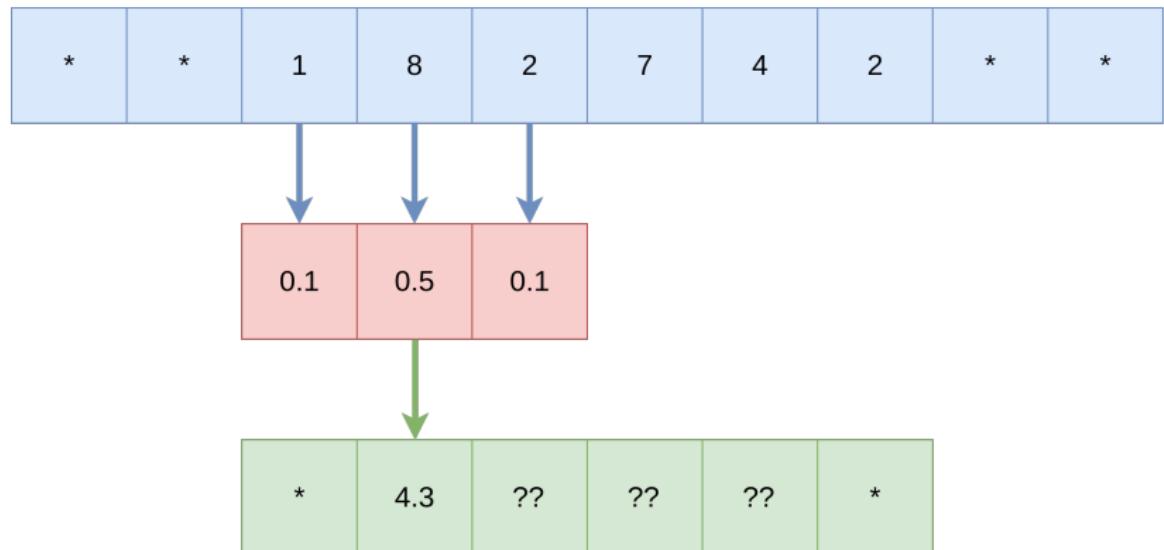


(b) Кот

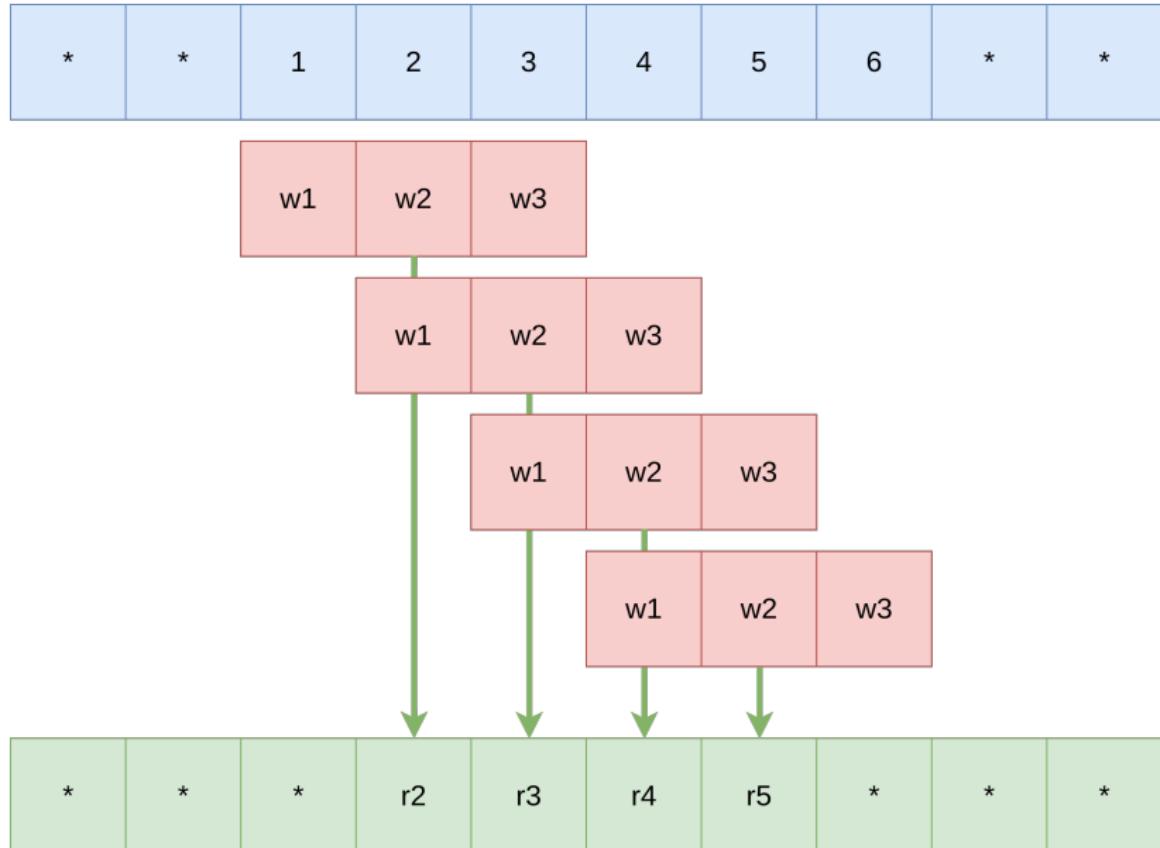
Одномерная свертка (convolution)

Определение

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



Одномерная свертка (convolution)



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	B	A
---	---	---	---	---	---	---

Циклический отступ

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

Двумерная свертка (чб картинки)

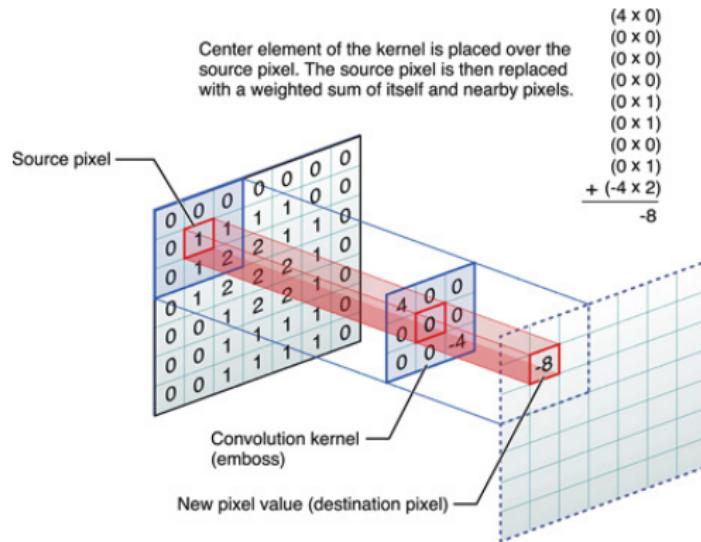


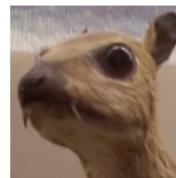
Figure: 2D convolution²

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

Примеры ядер

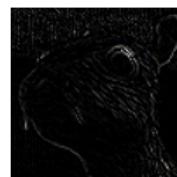
- ▶ Тождественное

0	0	0
0	1	0
0	0	0



- ▶ Детектор границ

0	1	0
1	-4	1
0	1	0



- ▶ Увеличение резкости

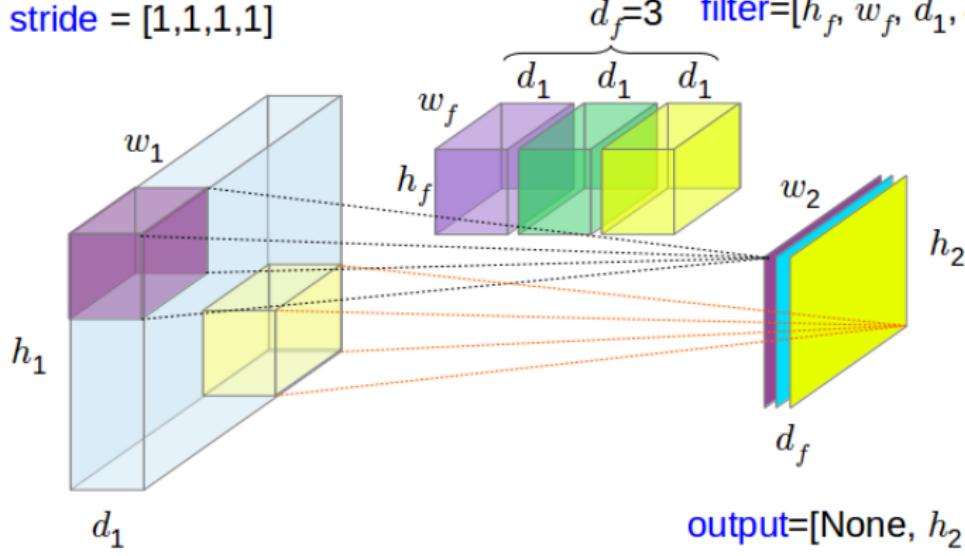
0	1	0
1	5	1
0	1	0



Свертка в нейронных сетях

padding = VALID
stride = [1,1,1,1]

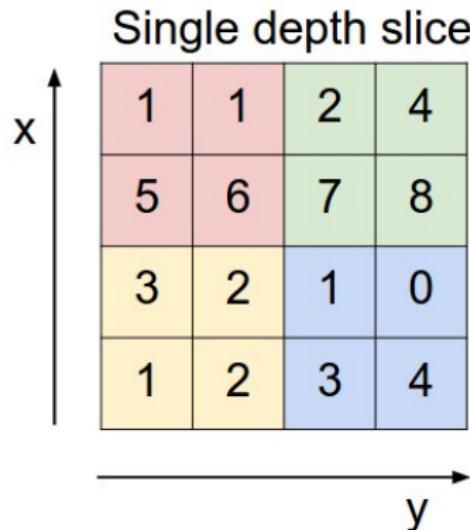
filter=[h_f , w_f , d_1 , d_f]



input=[None, h_1 , w_1 , d_1]

output=[None, h_2 , w_2 , d_f]

Pooling



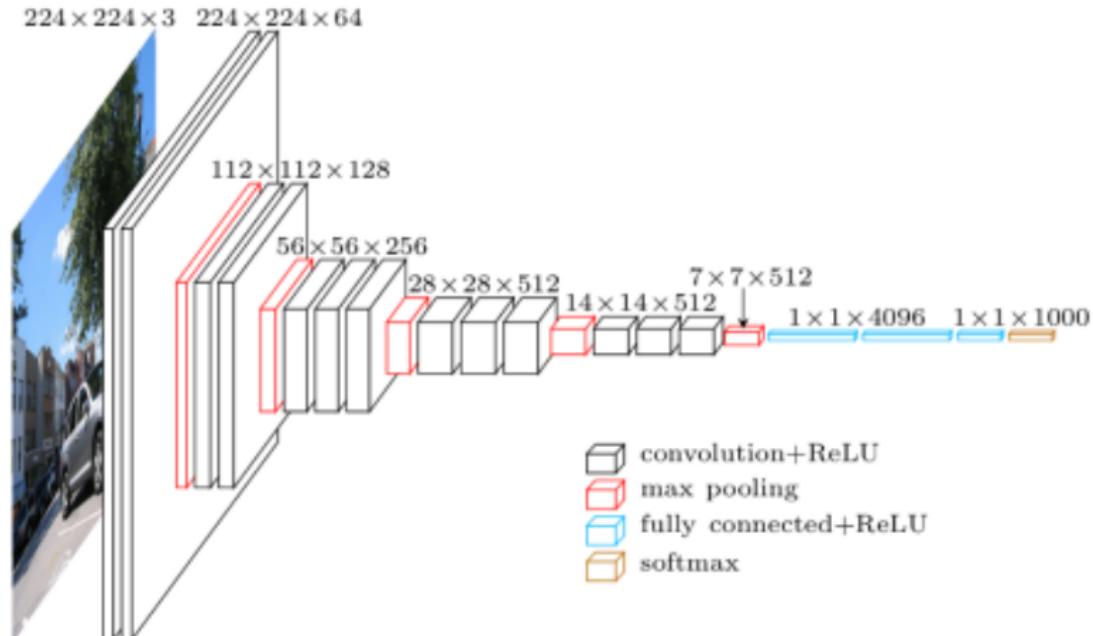
max pool with 2x2 filters
and stride 2



6	8
3	4

- ▶ 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

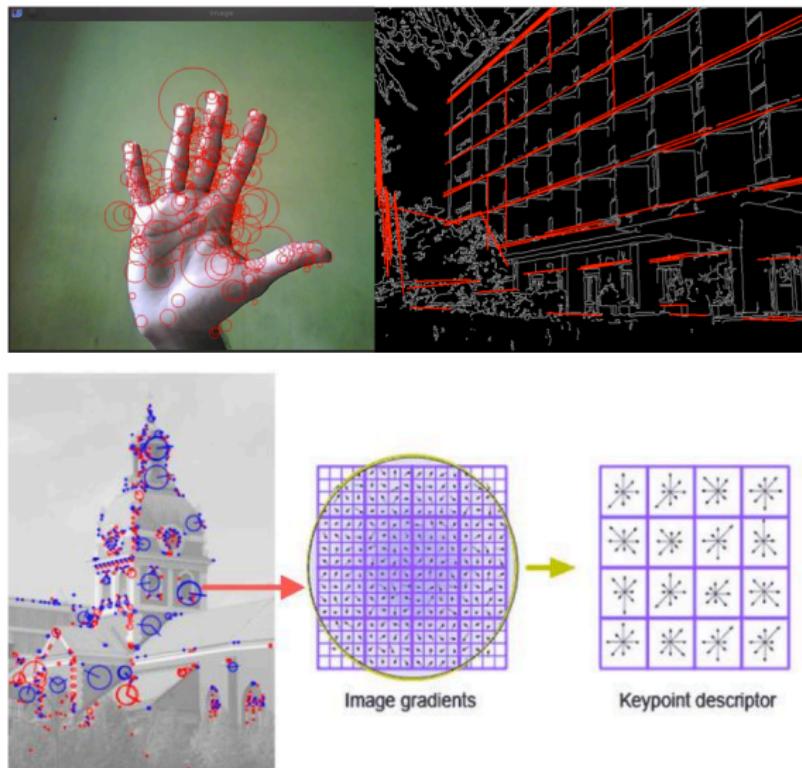


Figure: Классический подход к извлечению признаков

Извлечение признаков, история

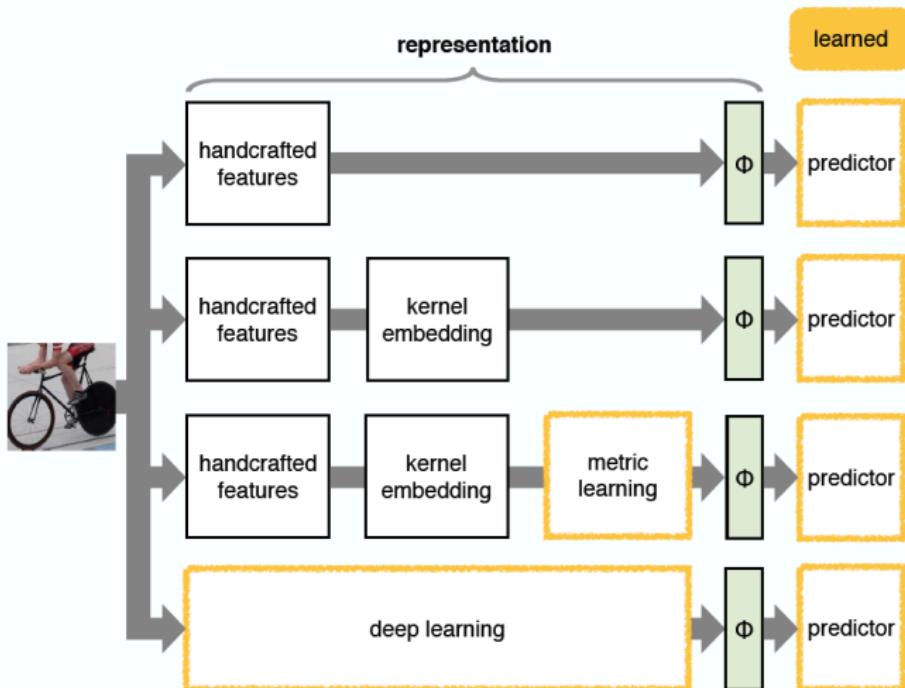
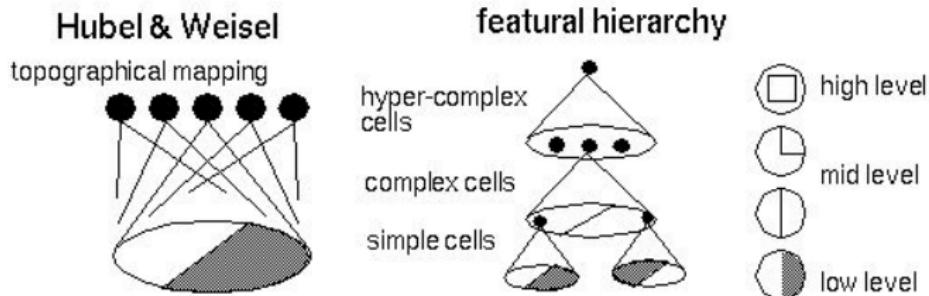


Figure: Глубинное обучение³

³Learning 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

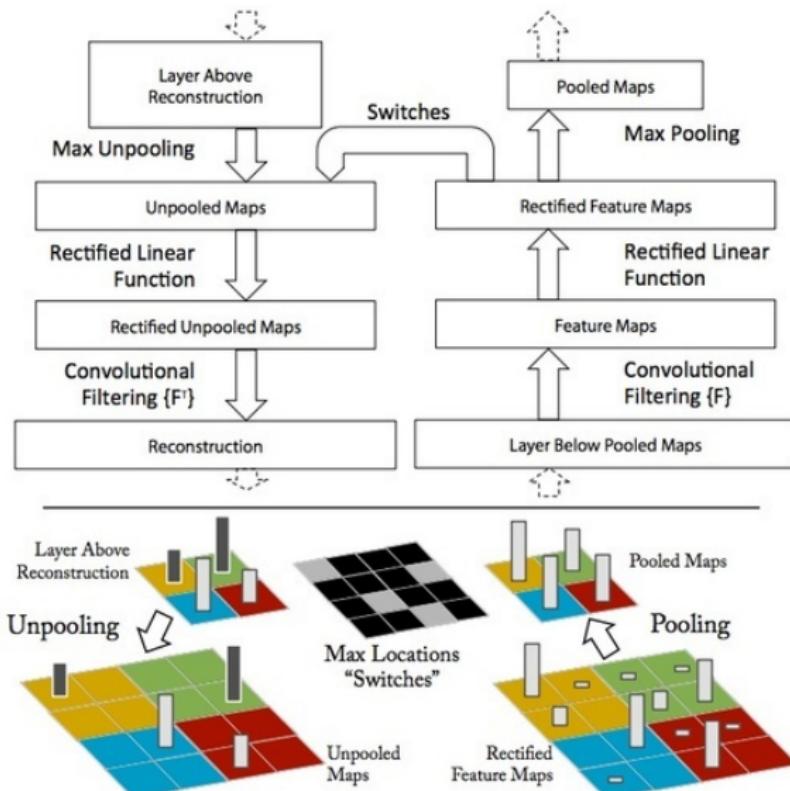


Figure: Схема deconvolution сети

Deconvolution сети

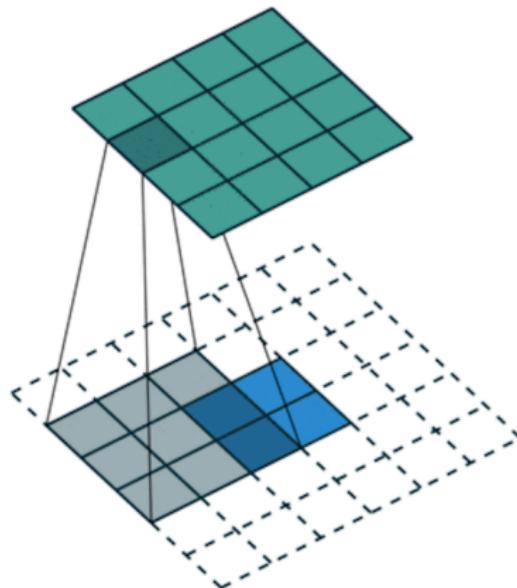


Figure: Convolution transposed⁴

⁴https://github.com/vdumoulin/conv_arithmetic

Deconvolution сети

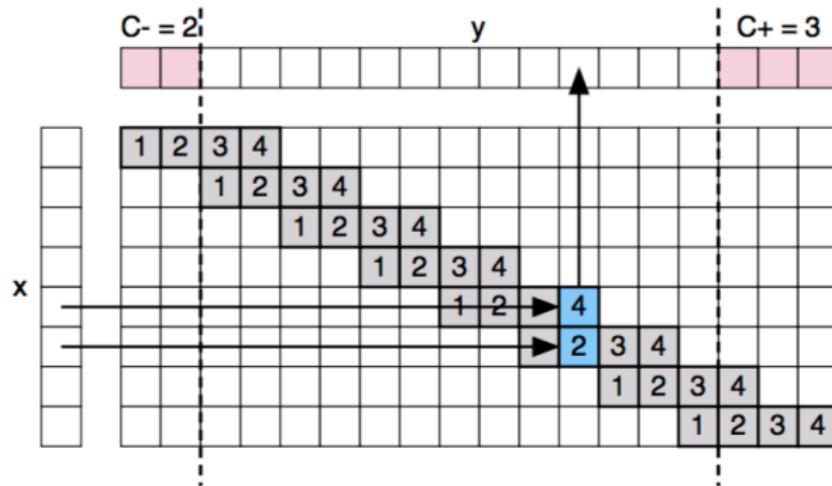


Figure: Convolution transposed⁵

⁵<http://www.vlfeat.org/matconvnet/matconvnet-manual.pdf>

Выучиваемые признаки

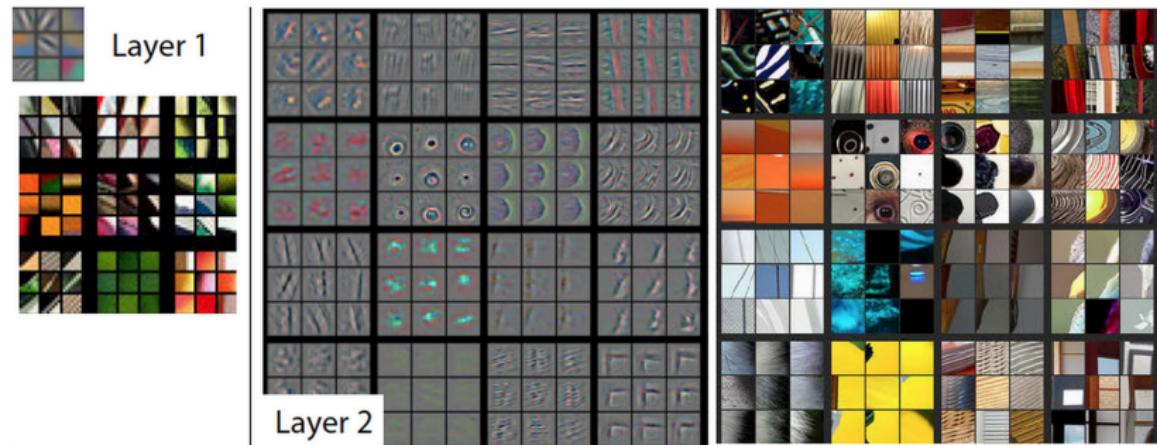


Figure: Visualizing and Understanding Convolutional Networks⁶

⁶Matthew D. Zeiler and Rob Fergus

Выучиваемые признаки

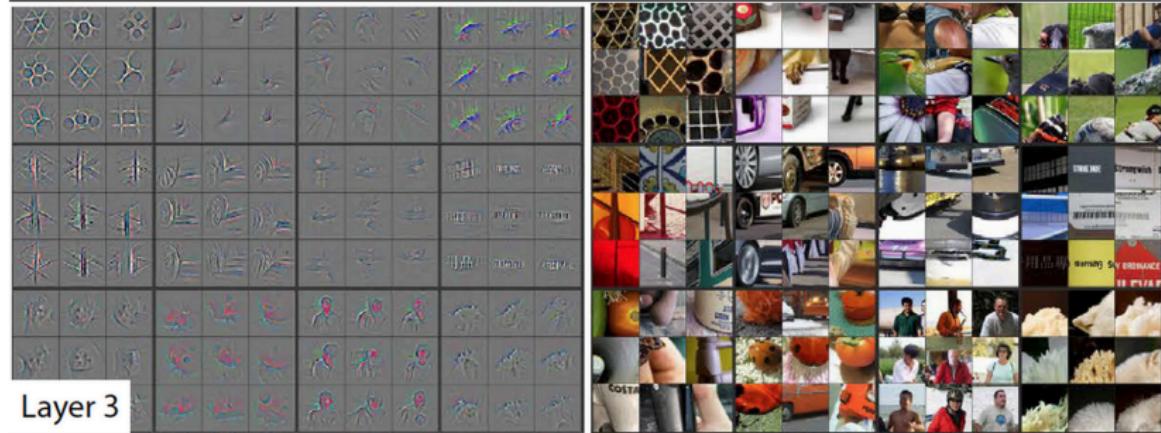


Figure: Visualizing and Understanding Convolutional Networks⁷

⁷Matthew D. Zeiler and Rob Fergus

Выучиваемые признаки

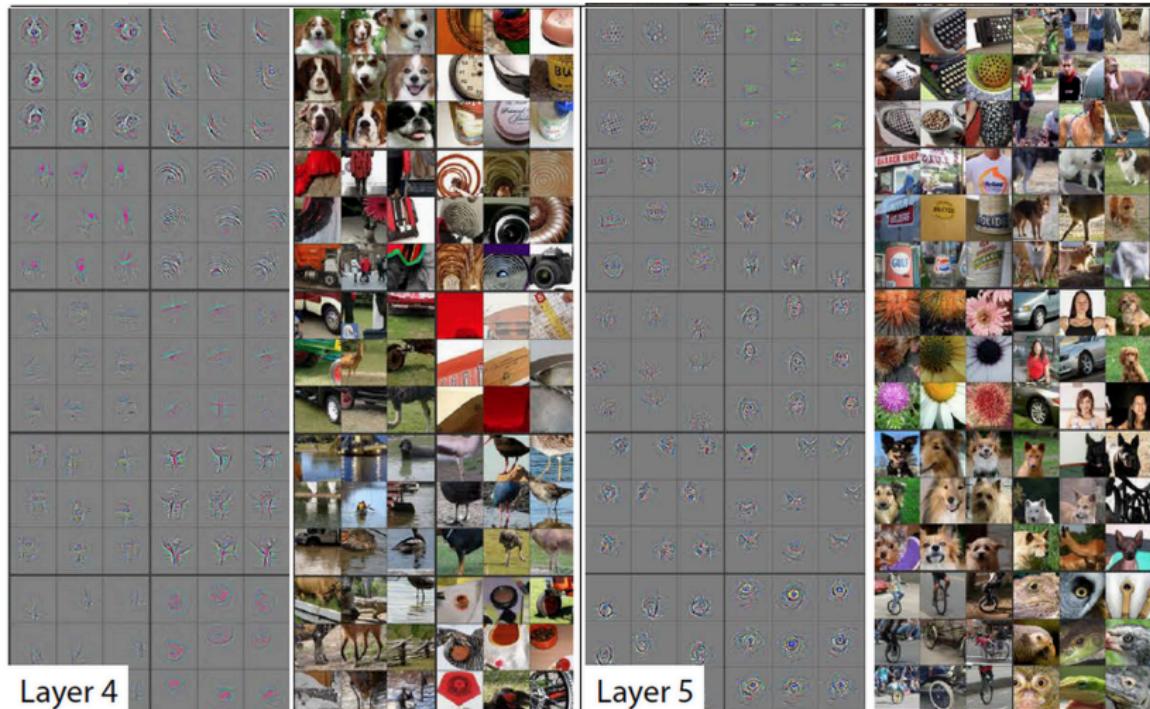


Figure: Visualizing and Understanding Convolutional Networks⁸

⁸Matthew D. Zeiler and Rob Fergus

Transfer learning

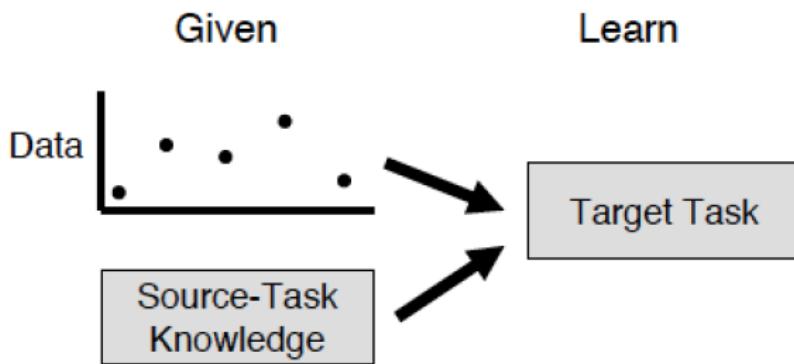


Figure: Модель решения задачи в рамках парадигмы трансфера знаний

Transfer learning

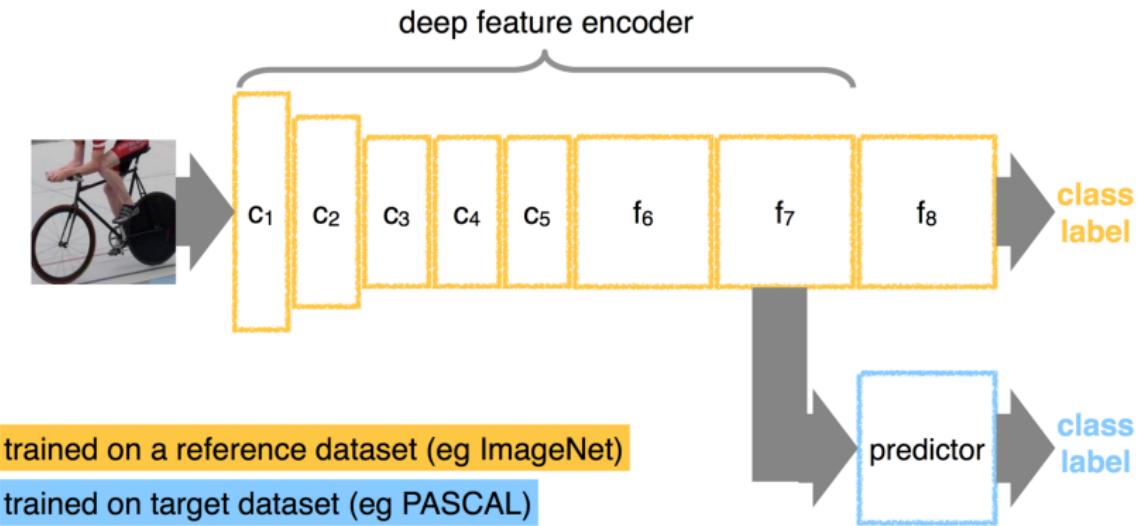


Figure: Трансфер между двумя глубинными сетями⁹

⁹Learning visual representations (Andrea Vedaldi)

Transfer learning¹⁰

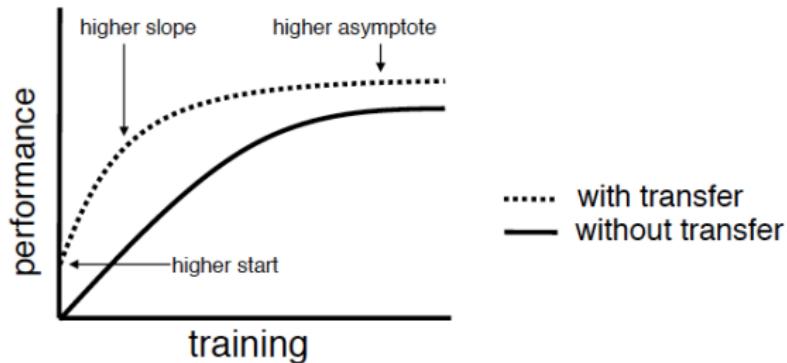


Figure: Цели трансфера знаний

- ▶ 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

¹⁰ftp:

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

Transfer learning

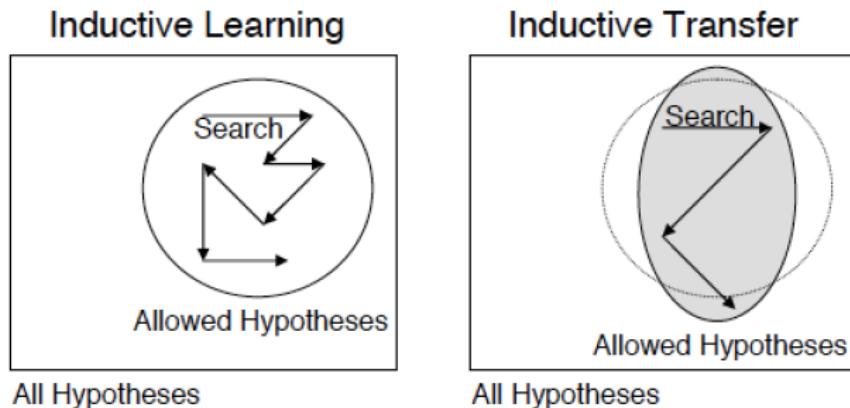


Figure: Трансфер знаний можно также рассматривать как некоторую регуляризацию, которая ограничивает пространство поиска до определенного набора допустимых и хороших гипотез

CNN для распознавания речи¹¹

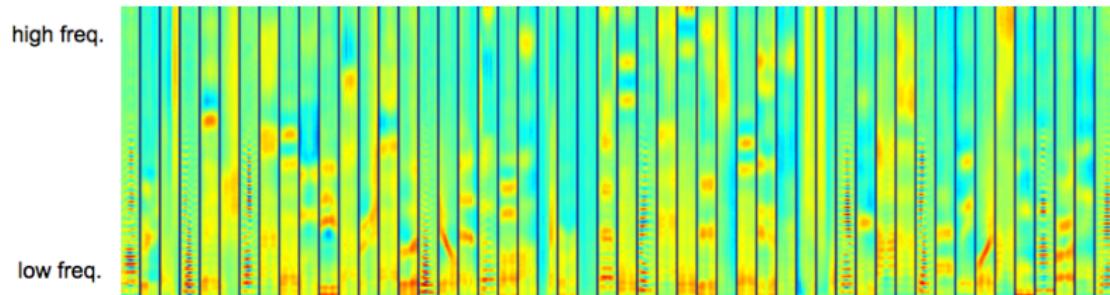


Figure: Спектрограмма голосового сигнала

¹¹<http://ai.stanford.edu/~ang/papers/nips09-AudioConvolutionalDBN.pdf>

CNN для текстов¹²¹³

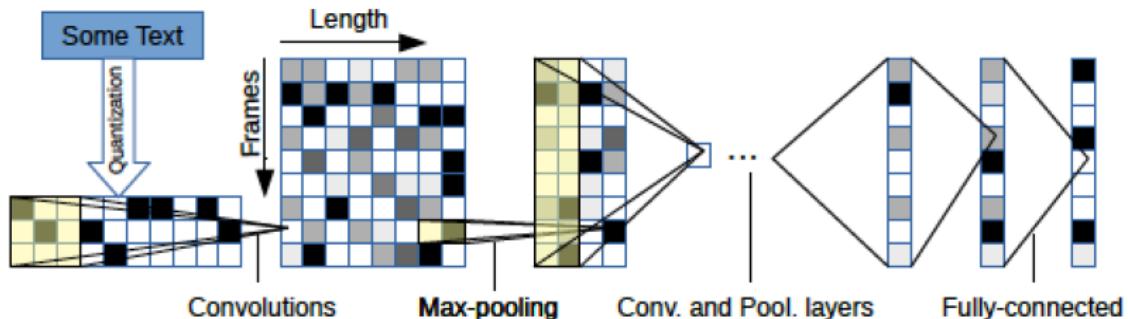


Figure: Обработка изображения представляющего текст

¹²<http://arxiv.org/pdf/1502.01710v1.pdf>

¹³http://nlp.csail.mit.edu/papers/Kalchbrenner_DCNN_ACL14

Вопросы

