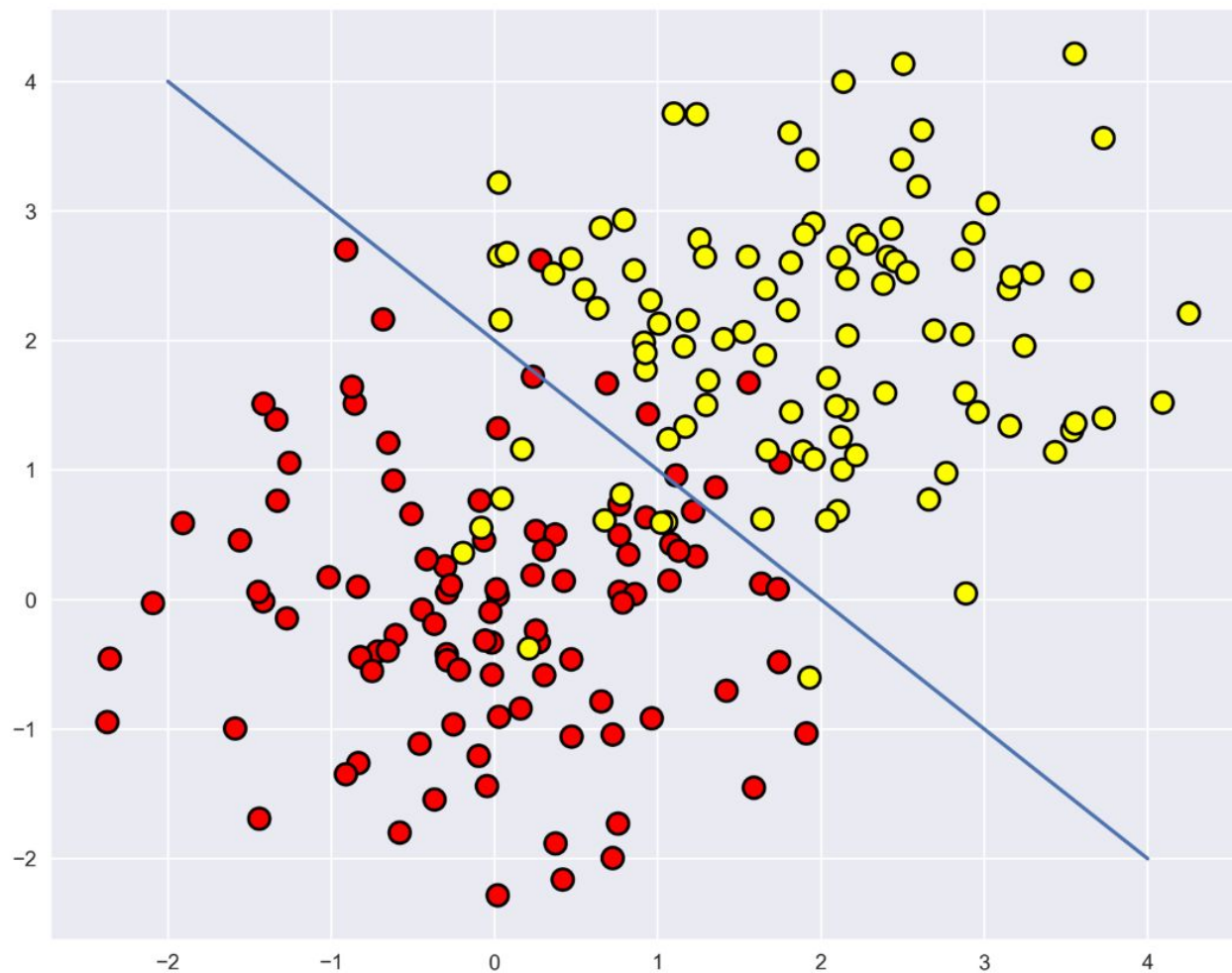


Deep learning in computer vision

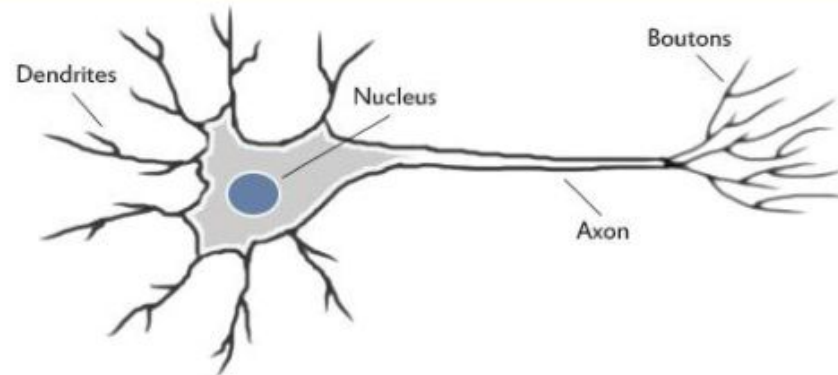


Попов Максим
Synesis, RnD

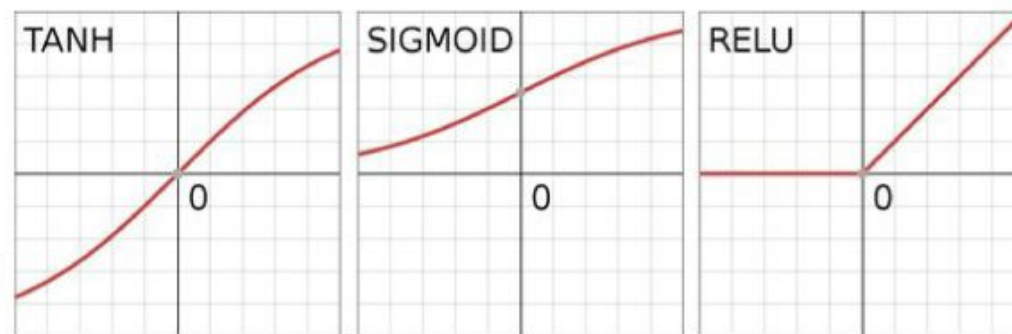
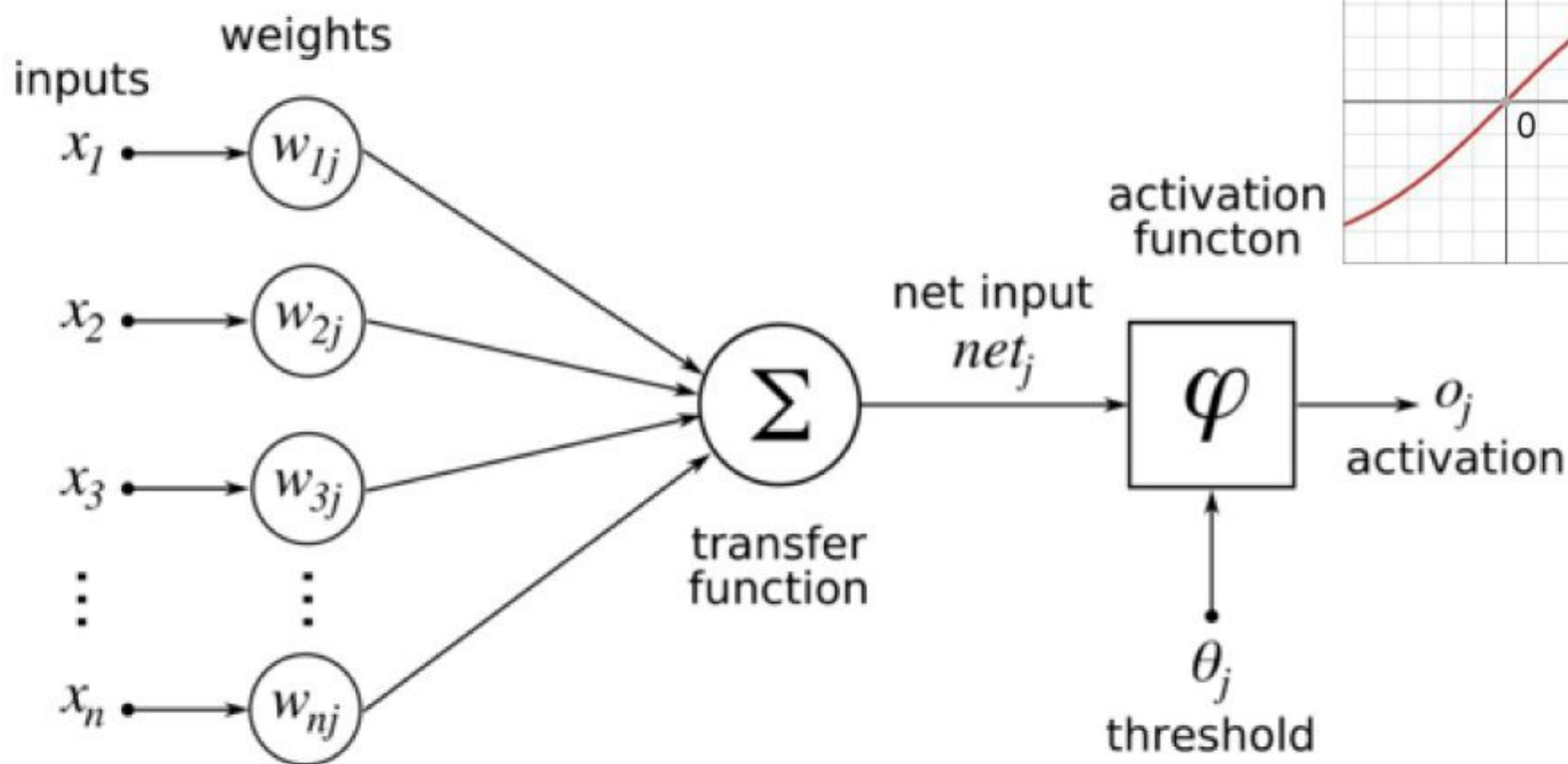
Задача классификации



Искусственный нейрон

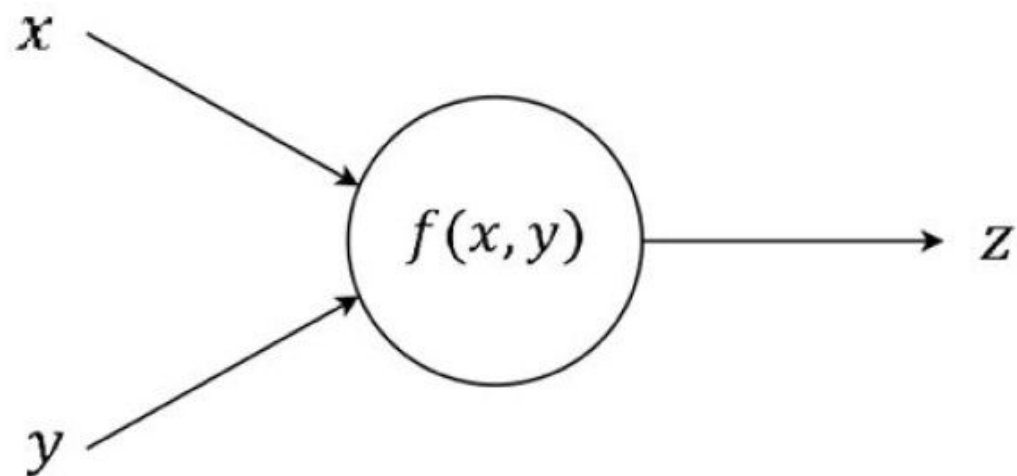


Искусственный нейрон — отдалённое подобие биологического. Базовый элемент искусственной нейронной сети

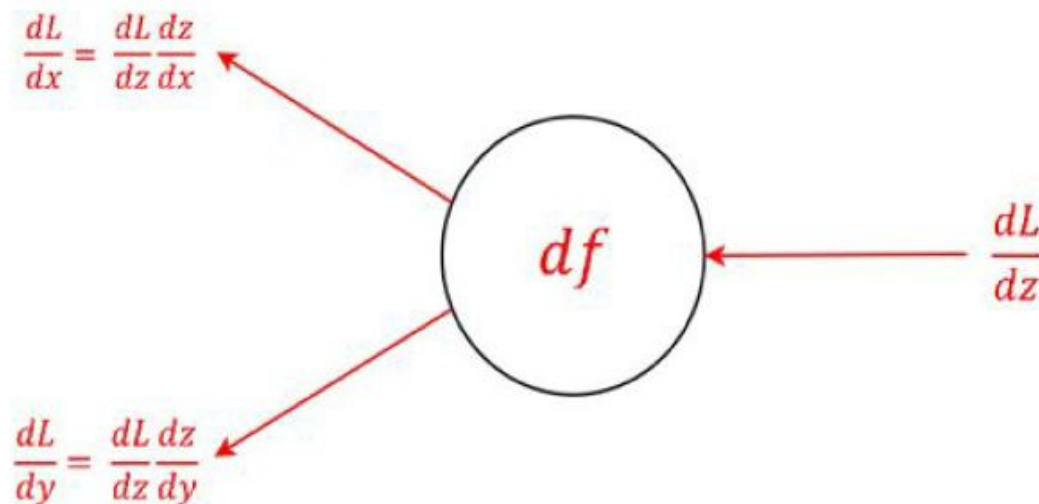


Обучение нейросети: backpropagation

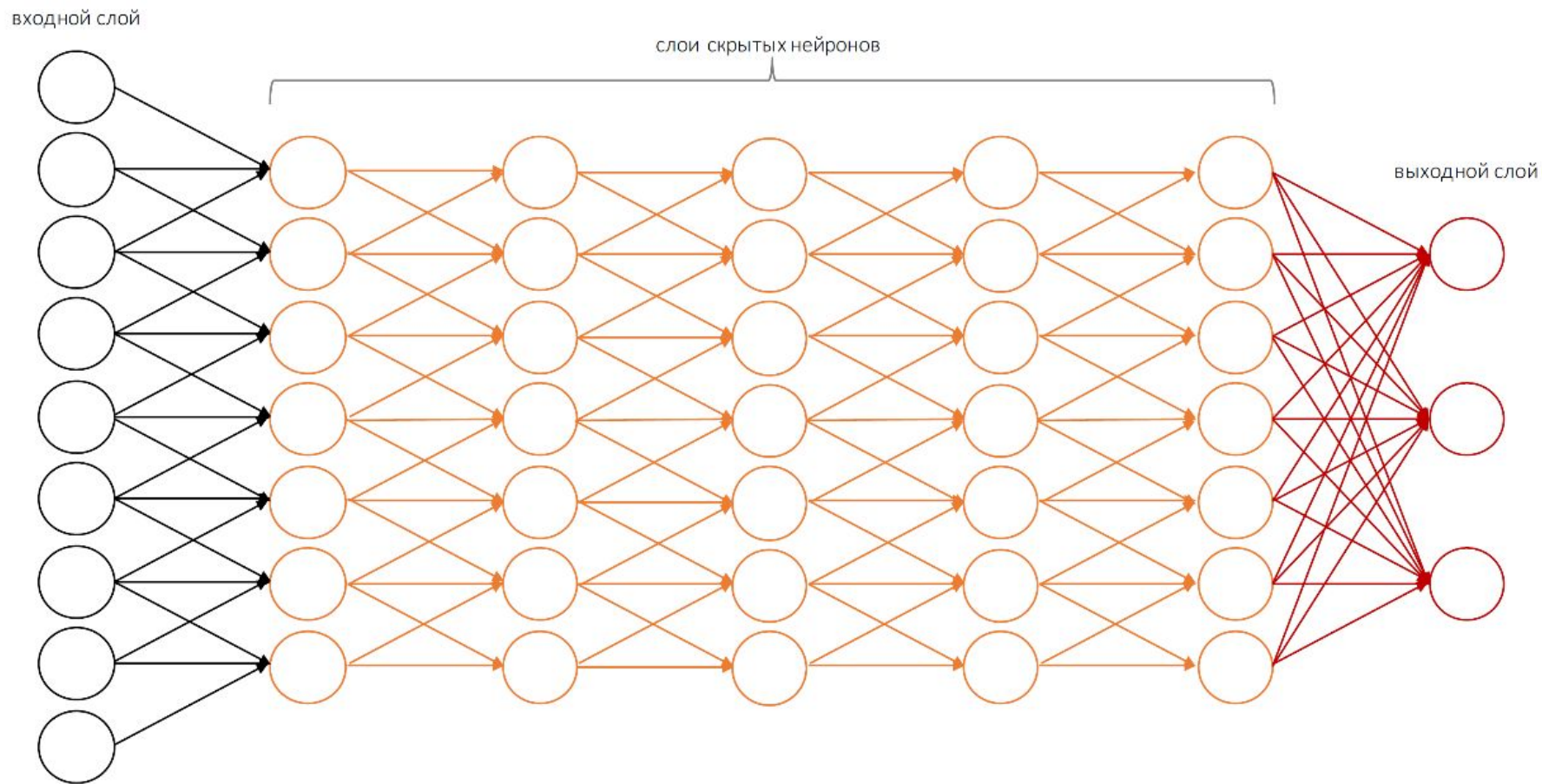
Forwardpass



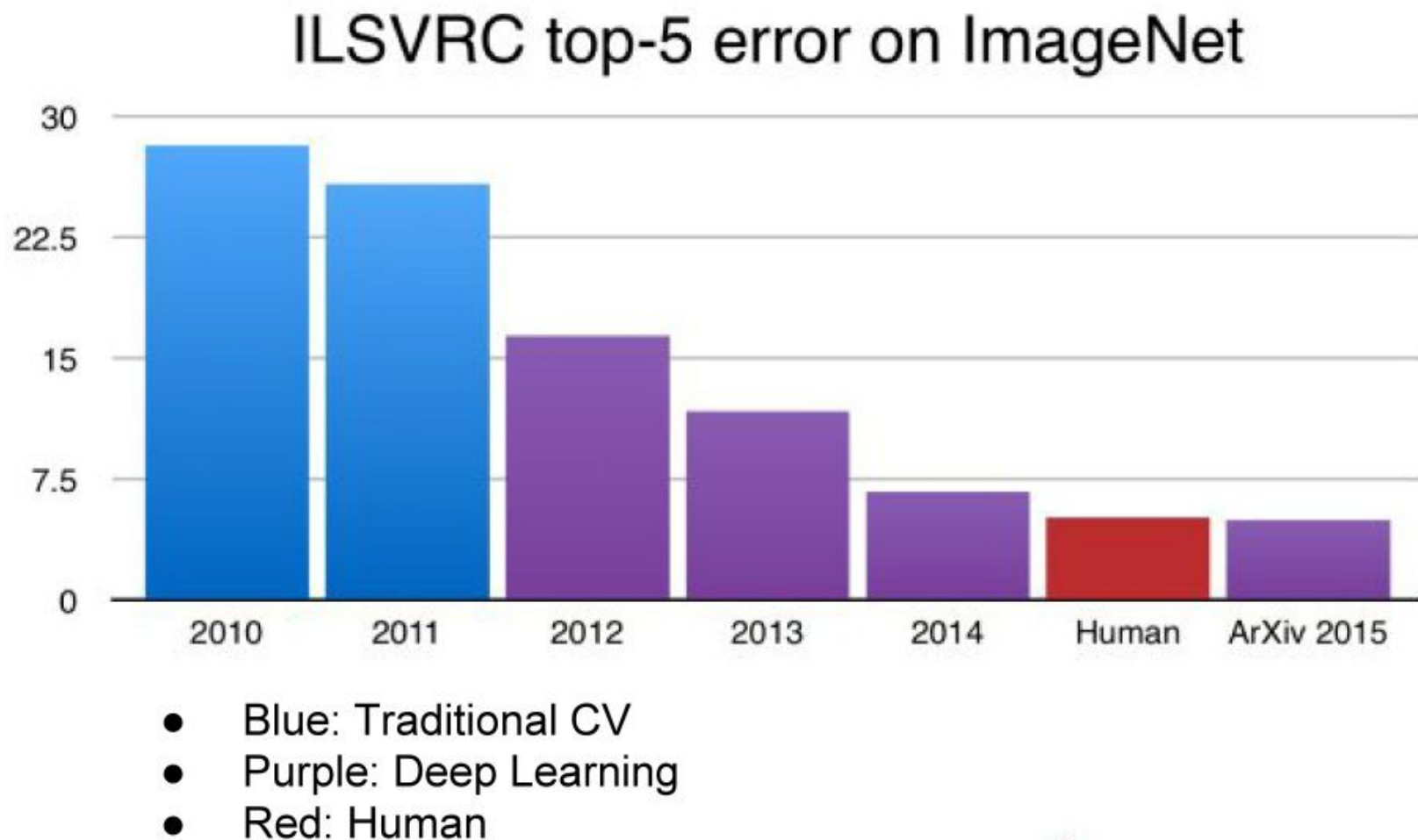
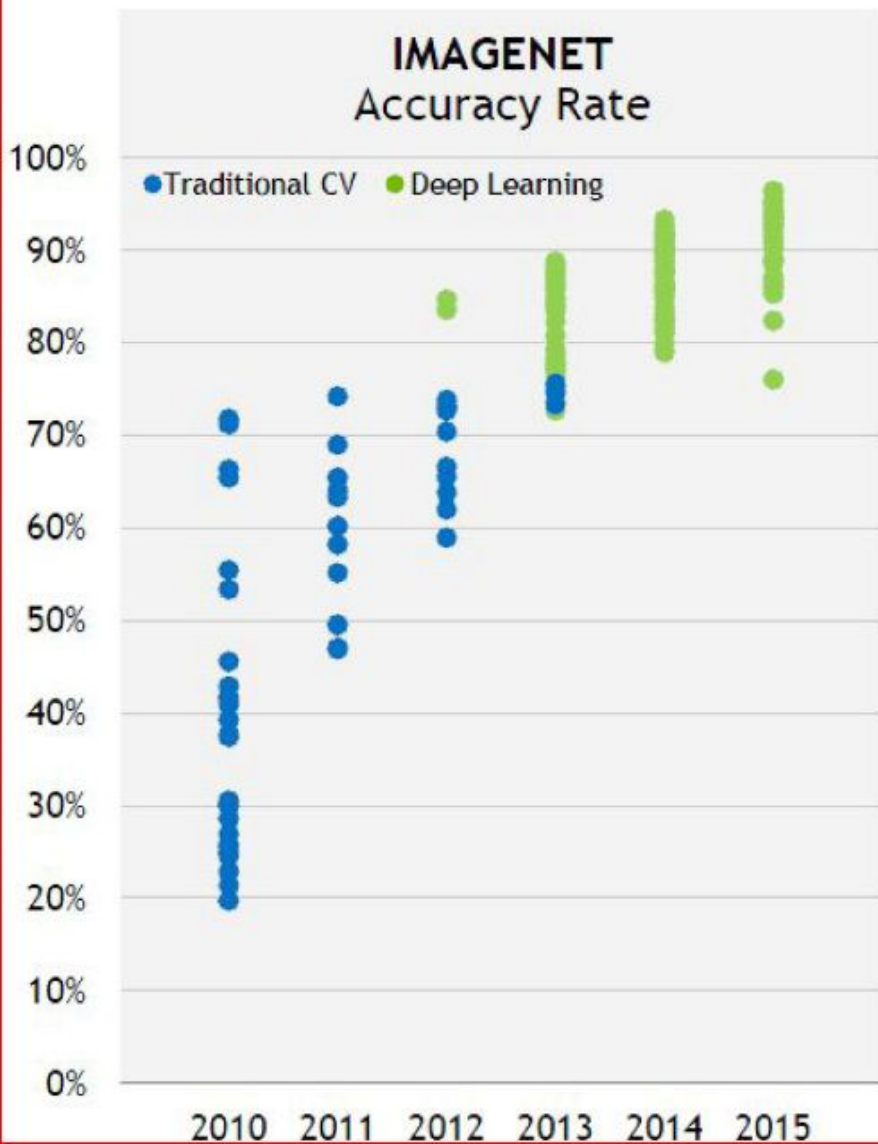
Backwardpass



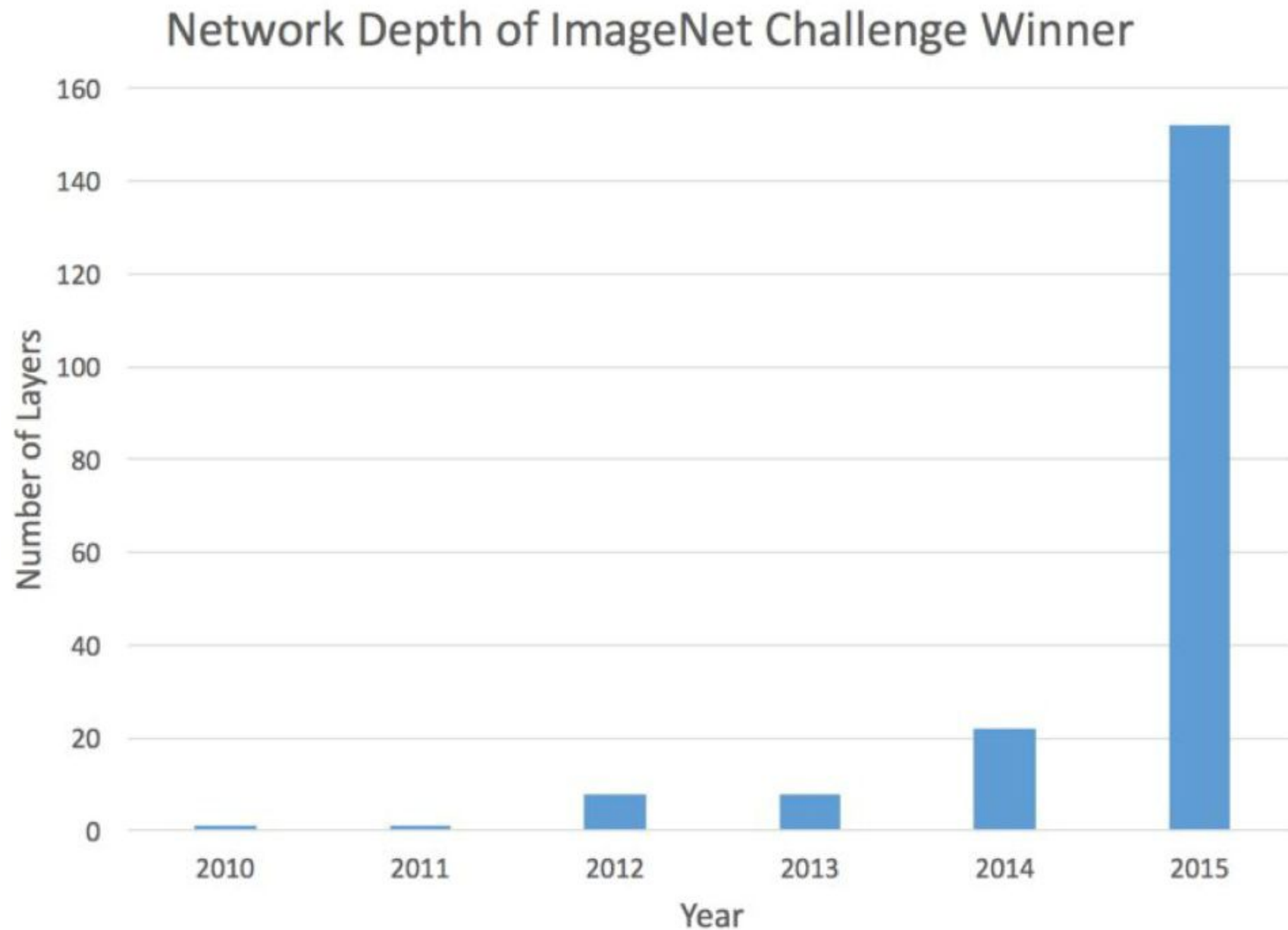
Искусственная нейросеть



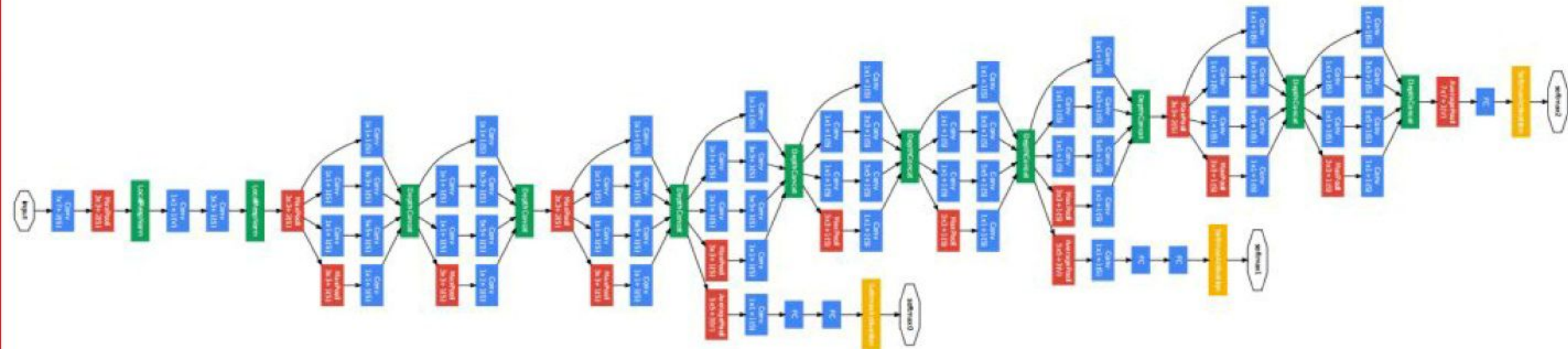
#1. Точность сетей растёт



#2. Сложность сетей растёт



#2. Сложность сетей растёт



Реальная нейросеть: GoogLeNet (2014)

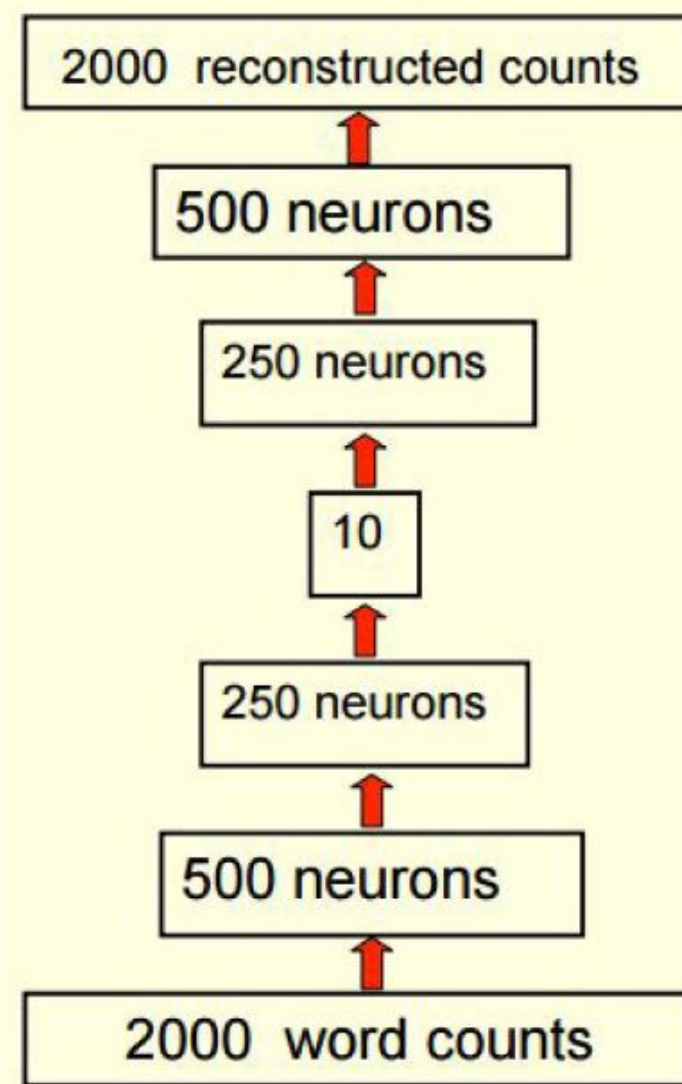
<http://cs.unc.edu/~wliu/papers/GoogLeNet.pdf>

Вариации FNN: Автоэнкодер (АЕ)

Учится создавать компактное описание входных данных.

Используется для уменьшения размерности и получения новых высокоуровневых признаков.

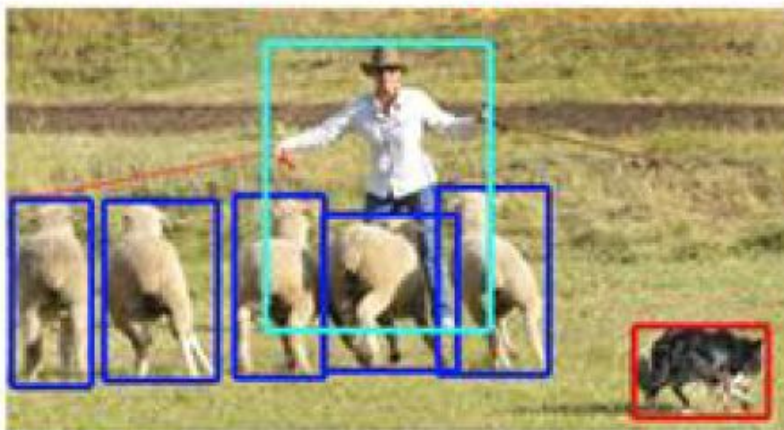
Может быть глубоким (многослойным).



Классические задачи для CNN



(a) classification



(b) detection

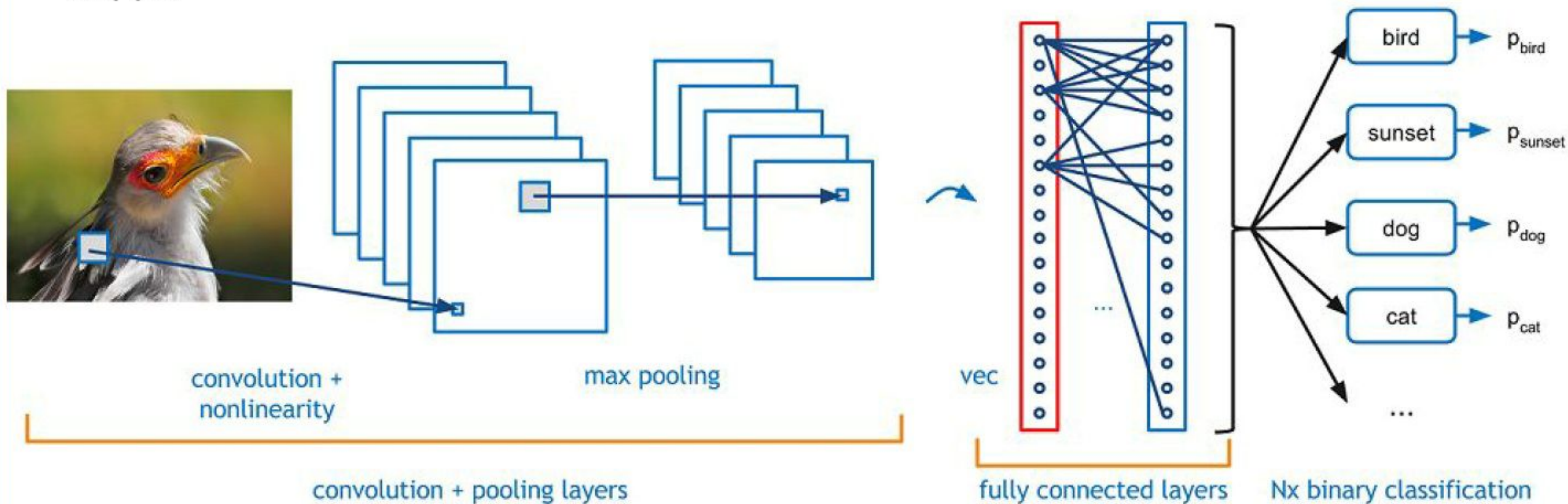


(c) segmentation

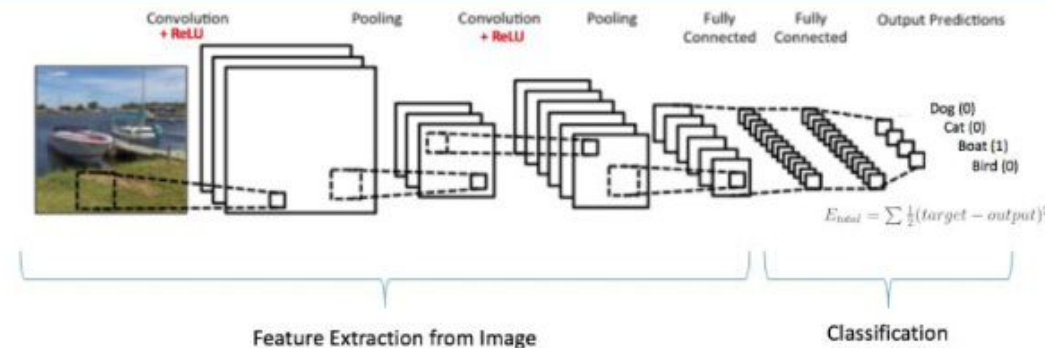
<https://research.facebook.com/blog/learning-to-segment/>

Свёрточная нейросеть: общий вид

Свёрточная нейросеть (CNN) — это Feed-Forward сеть специального вида:



Состав CNN



- **Свёрточные слои:** каждая плоскость в свёрточном слое — это один нейрон, реализующий операцию свёртки (convolution) и являющийся матричным фильтром небольшого размера (например, 5x5).
- **Слои субдискретизации** (subsampling, spatial pooling): уменьшают размер изображения (например, в 2 раза).
- **Полносвязные слои** (MLP) на выходе модели (используются для классификации).

Визуализация операции свёртки

Знакомые по фотошопу фильтры blur, emboss, sharpen и другие — это именно матричные фильтры.

	1	1	2	
	2	2	1	
	1	5	0	

input



2	3	0
0	1	3
3	0	3

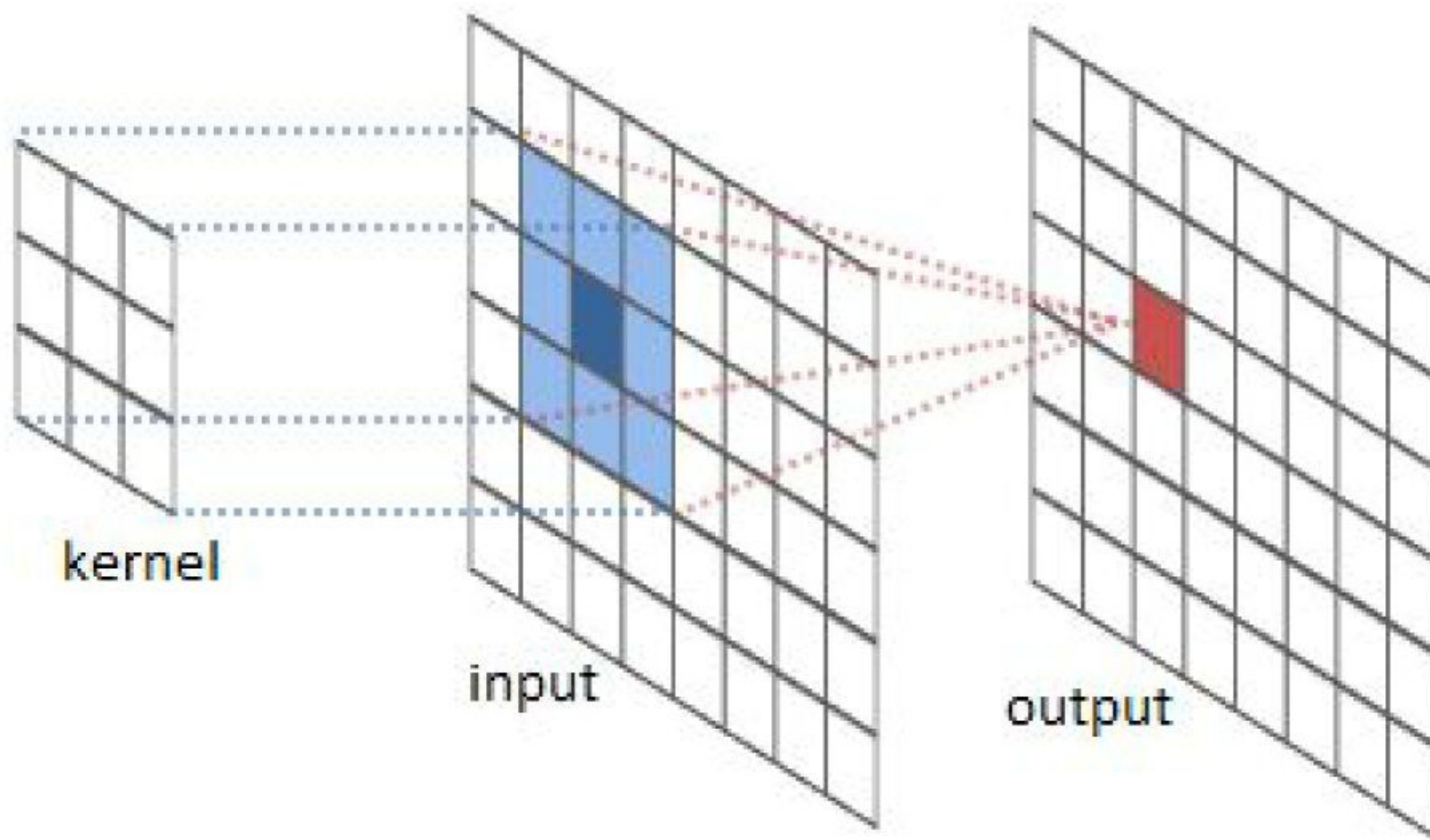
kernel



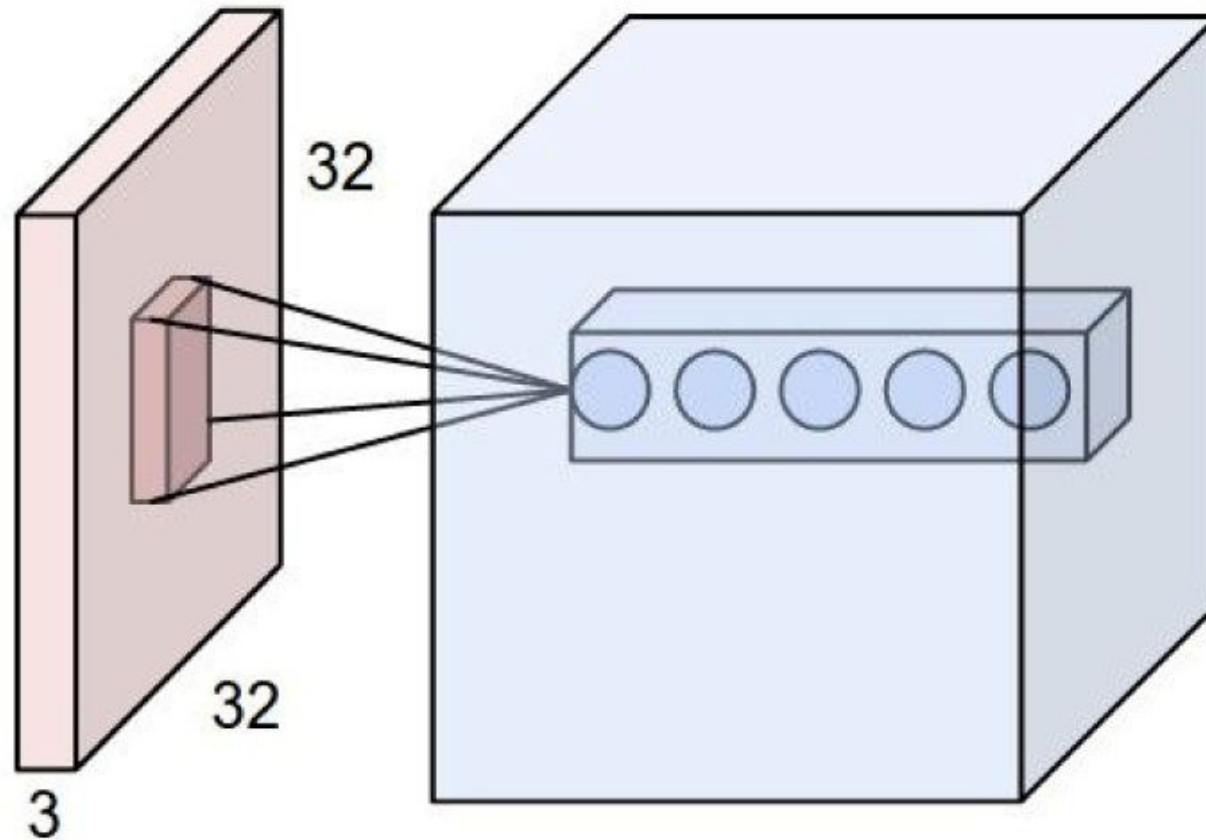
	1	1	2	
	2	13	1	
	1	5	0	

output

Визуализация операции свёртки

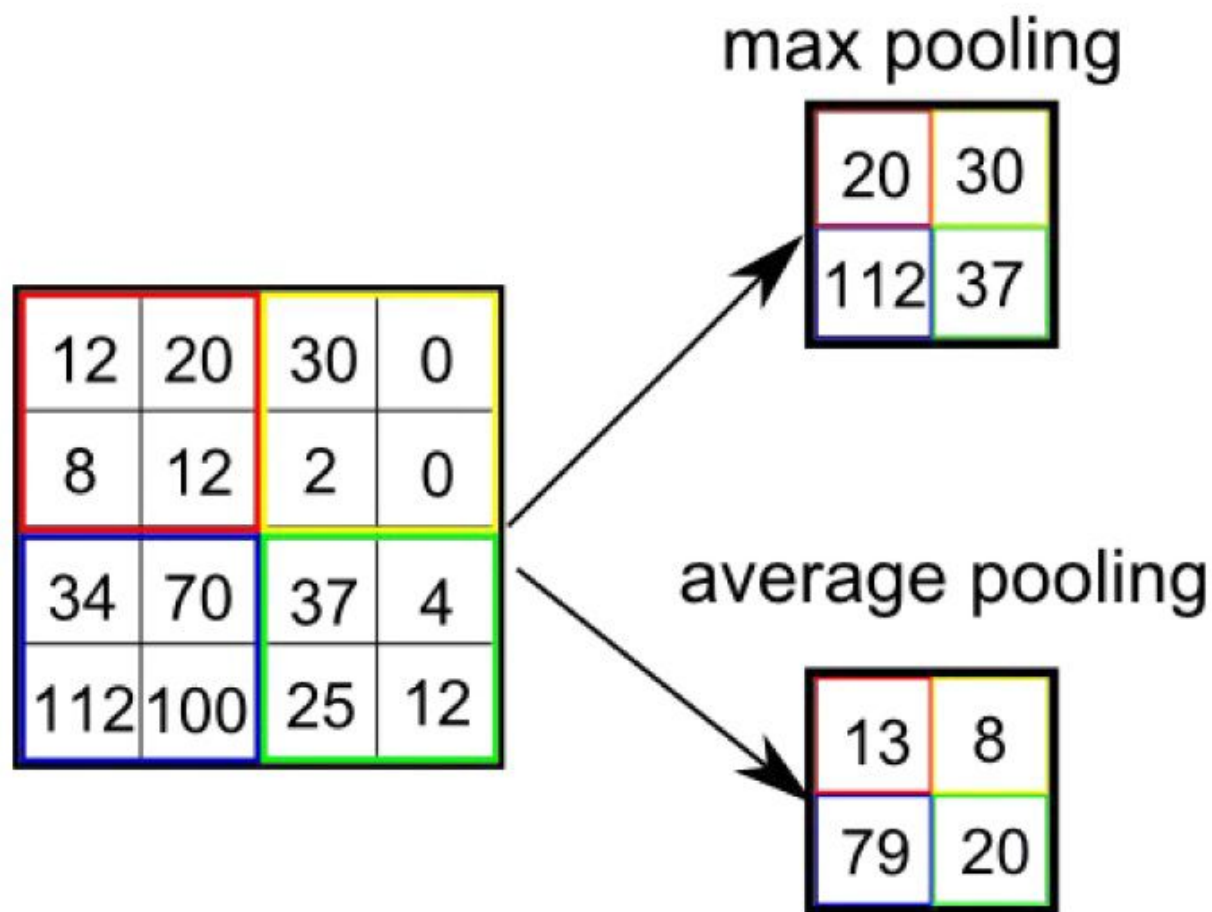


CNN: Свёрточный слой (5 нейронов)

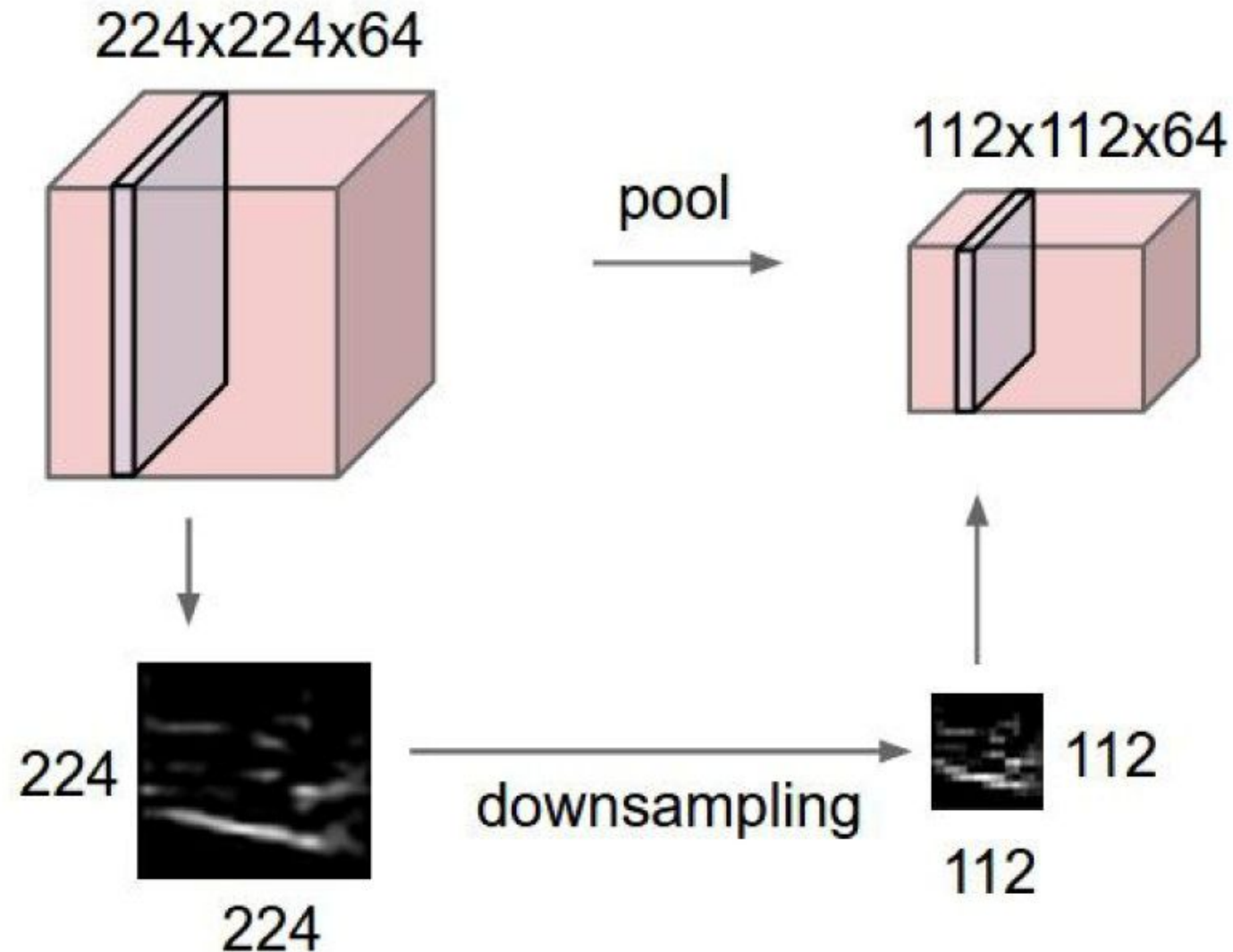


Веса нейронов — это коэффициенты ядра свёртки. Каждая “обучаемая” свёртка выделяет одинаковые локальные признаки во всех частях изображения.

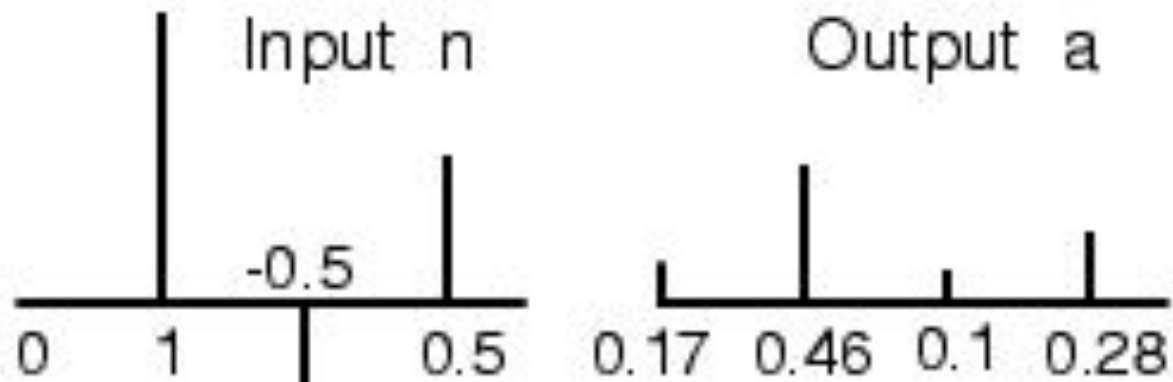
Операция pooling (max pool, avg pool)



CNN: Pooling слой (downsampling)

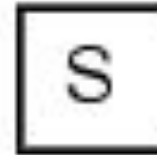


Softmax function



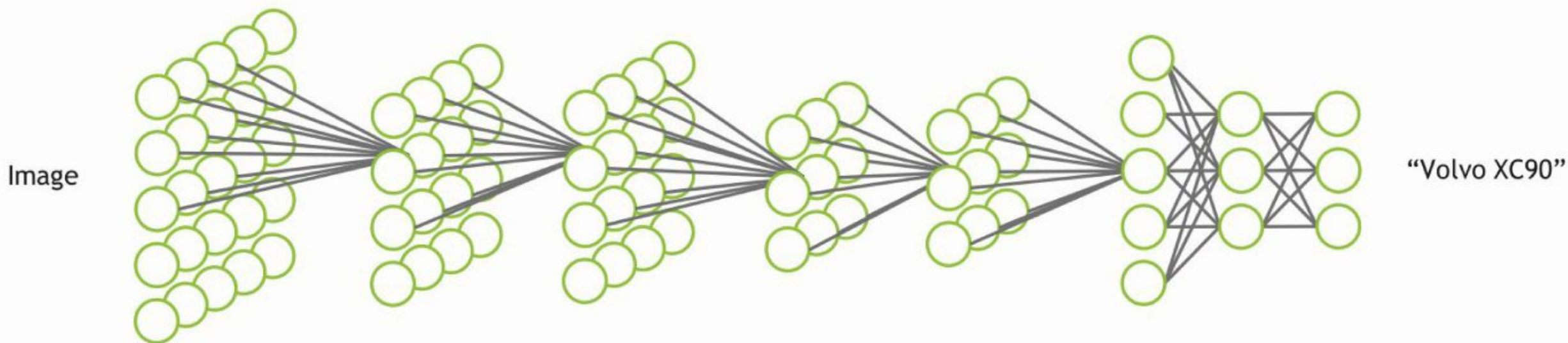
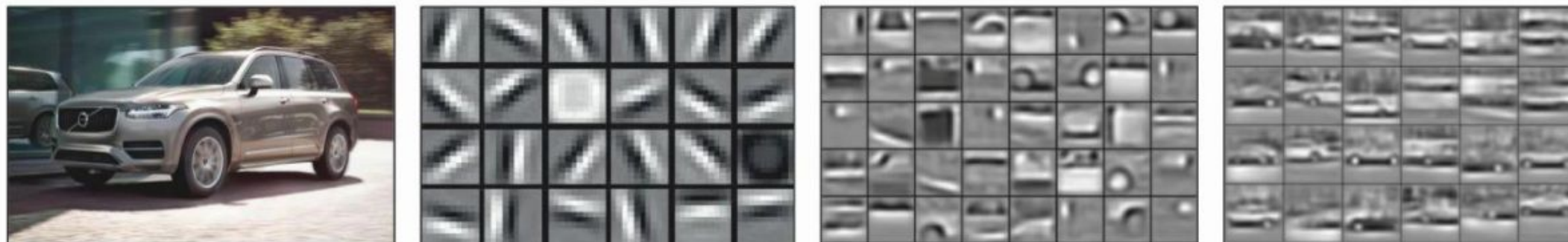
$$a = \text{softmax}(n)$$

Softmax Transfer Function



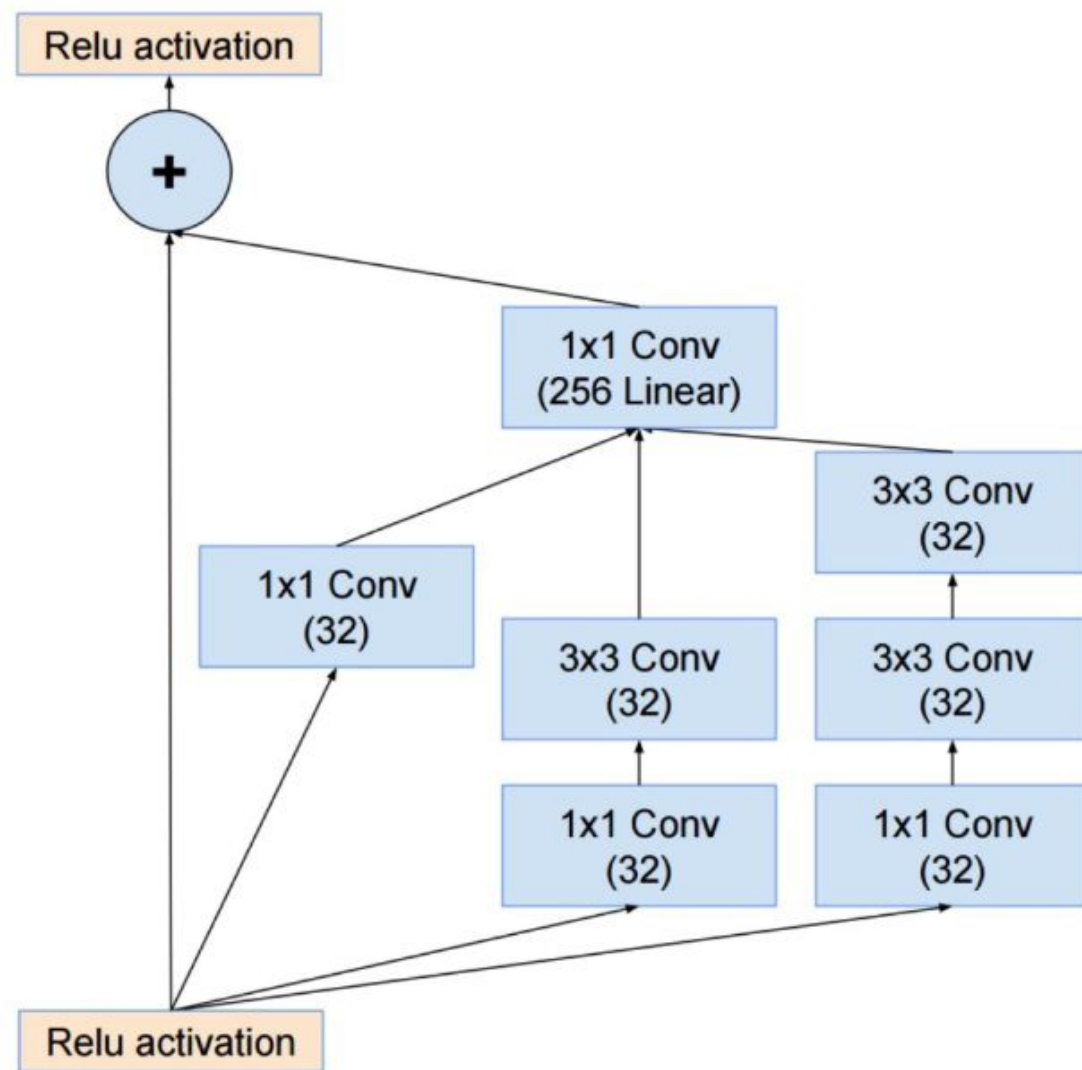
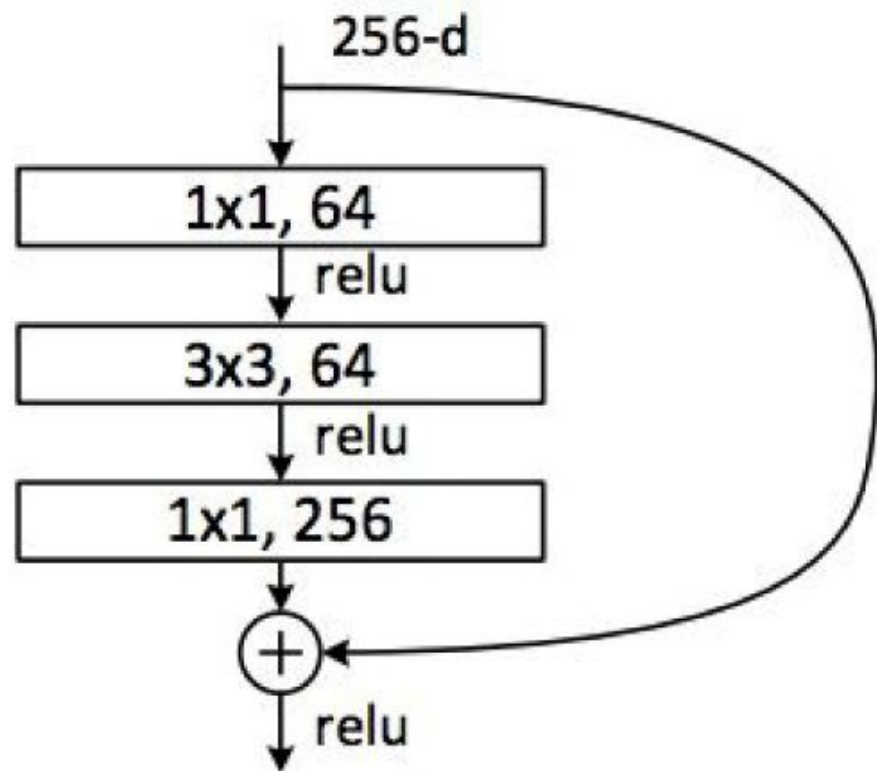
Свёрточная нейросеть

Свёрточные слои учат иерархические признаки для изображений, а spatial pooling даёт некоторую инвариантность к перемещениям.



Современные архитектуры

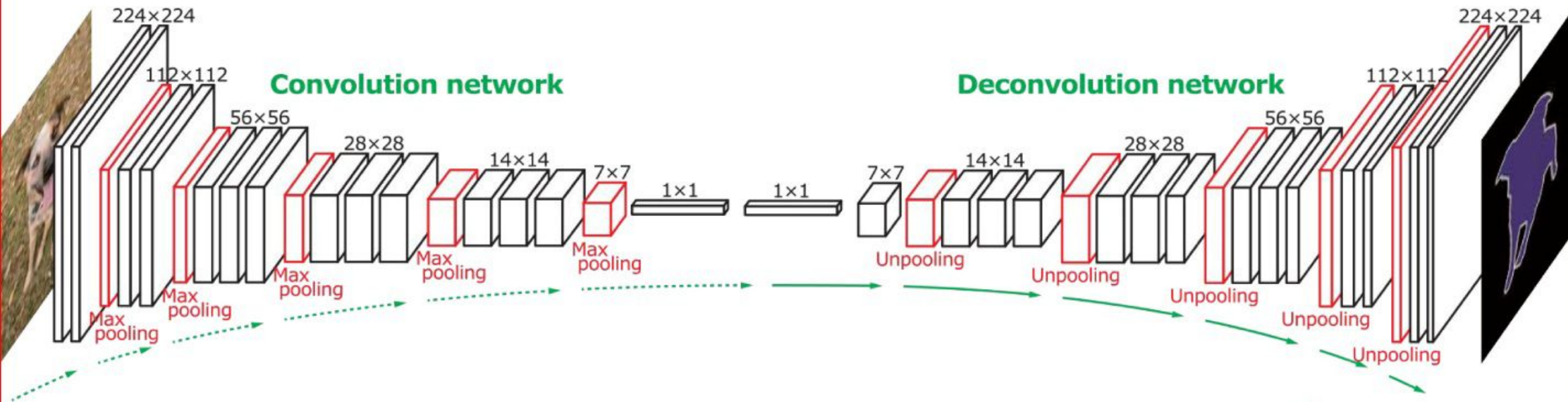
Inception, ResNet и другие современные архитектуры содержат специальные блоки слоёв.



Deconvolution networks

Правильнее называть это Transposed convolution, а не Deconvolution (это слово уже занято в цифровой обработке сигналов для обратной операции).

По сути, реализован обучаемый upsampling.



Пример кода на Keras

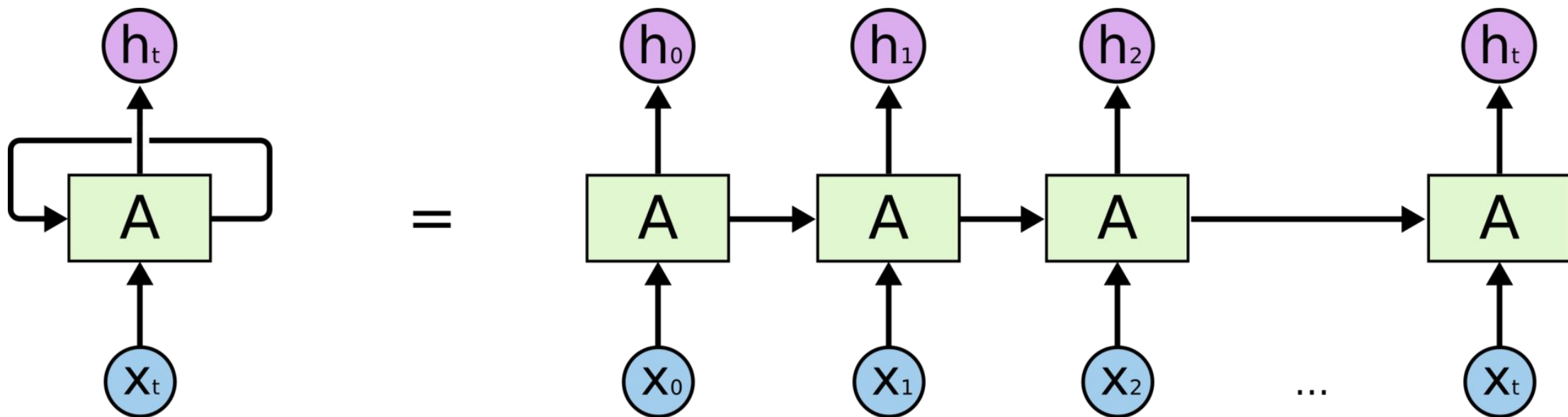
```
Model = keras.Sequential([  
    keras.layers.Flatten(input_shape = (28,28)),  
    keras.layers.Dense(128, activation = "relu"),  
    keras.layers.Dense(10, activation = "softmax")  
])
```

```
model.compile(optimizer="sgd",  
    loss = "categorical_crossentropy",  
    metrics = ["accuracy"])
```

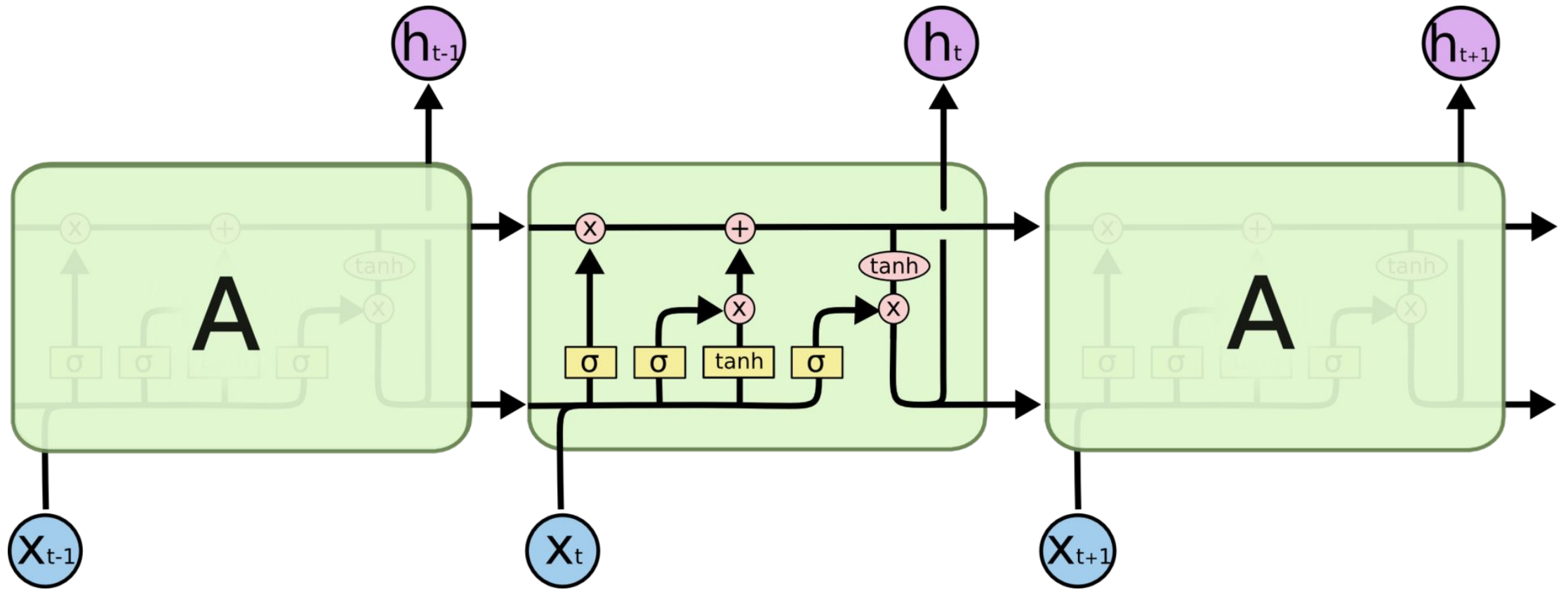
```
model.fit(train_images, train_labels, epoch = 10)
```

```
predict = model.predict(test_images)
```

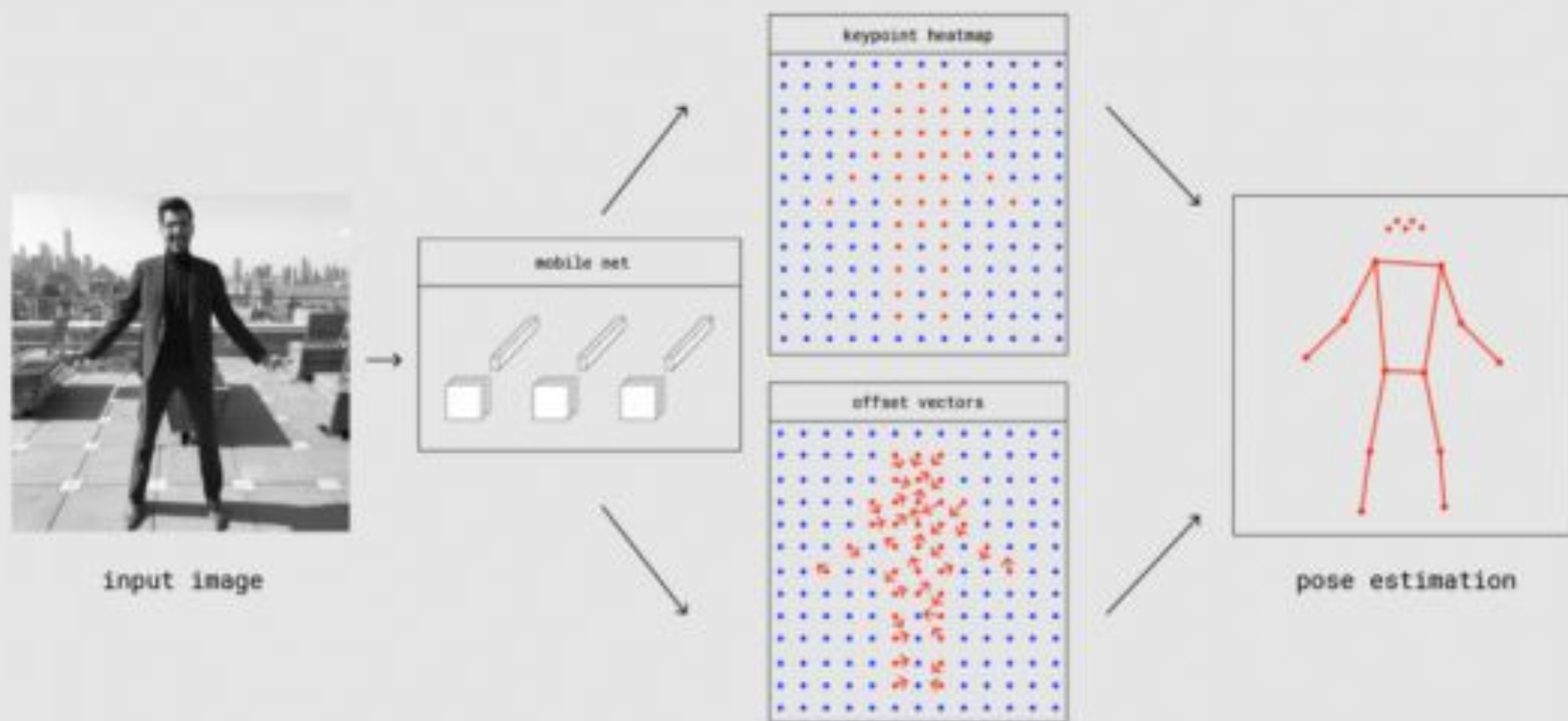
Рекуррентные нейронные сети



Сети LSTM (Long short-term memory)



Single-Pose Detection Algorithm



input image

mobile net

keypoint heatmap

offset vectors

pose estimation

PoseNet model