

Использование нейронных сетей для анализа видео

Спасёнов Алексей
Старший программист-исследователь
Команда Компьютерного зрения



Содержание лекции

1. Основы компьютерного зрения
2. Способы представления изображений
3. Введение в машинное обучение
4. Нейронные сети в компьютерном зрении
5. Описание задания



ОСНОВЫ КОМПЬЮТЕРНОГО ЗРЕНИЯ



Задача компьютерного зрения



Мы видим



Компьютер видит

Задача компьютерного зрения

Классификация Классификация с локализацией Детекция объектов Сегментация экземпляров

CAT CAT CAT, DOG, DUCK CAT, DOG, DUCK

Один объект Много объектов

Buildings Forest Mountains

Glacier Sea Street

Классификация сцен

3D

3D reconstruction of a person's head wearing a yellow hat, showing multiple views and a depth map.

Tracking

Tracking

Legend:

1		
2		3
4		5
	6	
7		8

Задача компьютерного зрения



(a) Illumination



(b) Deformation



(c) Scale, Viewpoint



(d) Pose, Occlusion



(e) Clutter, Occlusion



(f) Blur



(g) Motion



(h) Small Objects,
Low Resolution



(i) Different instances of the “chair” category

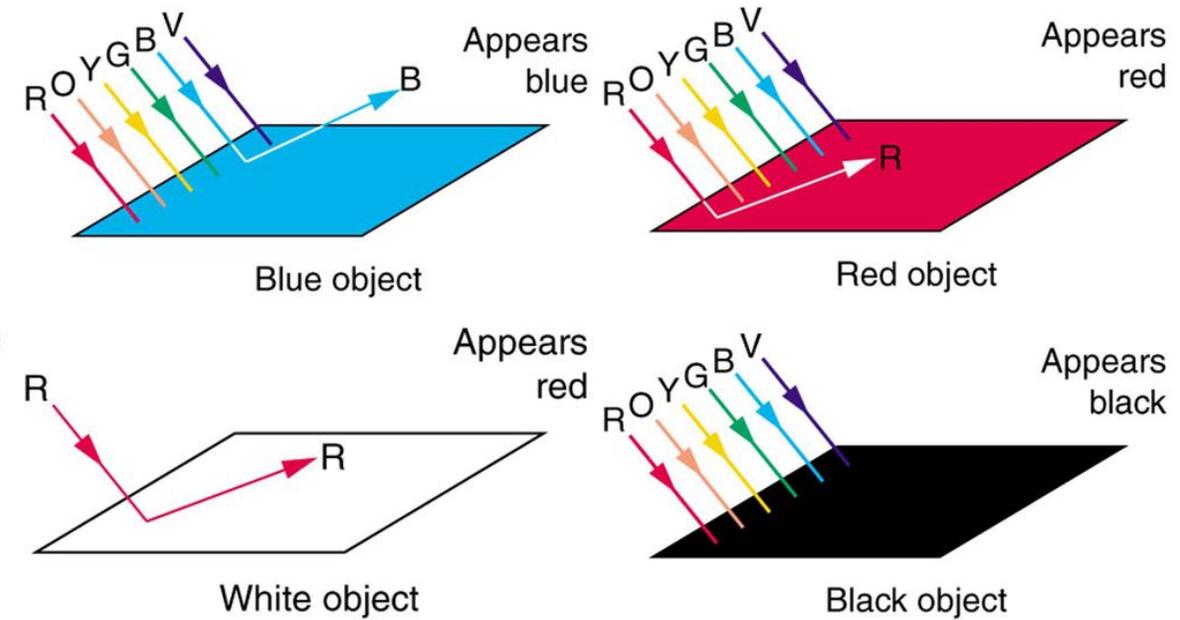
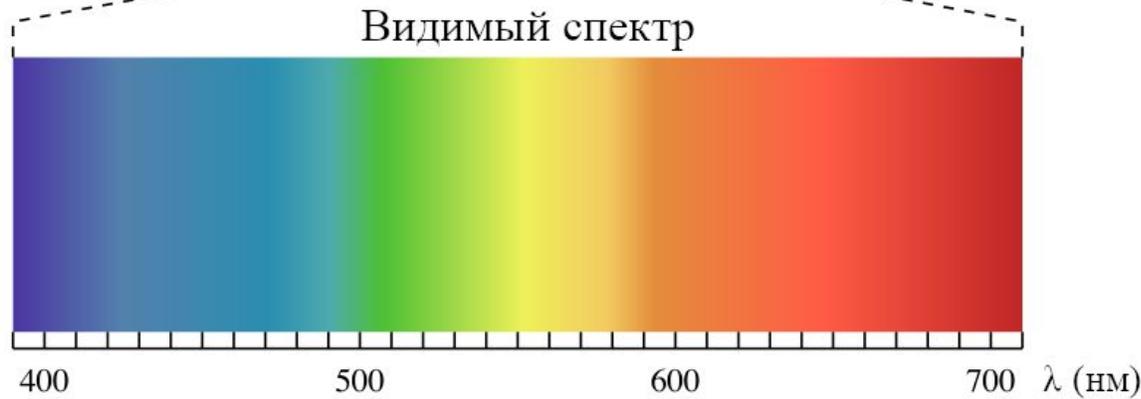
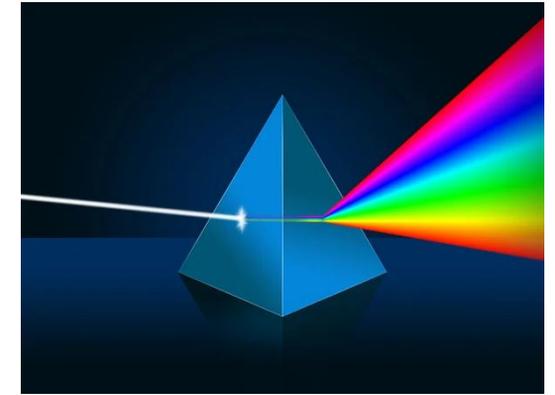
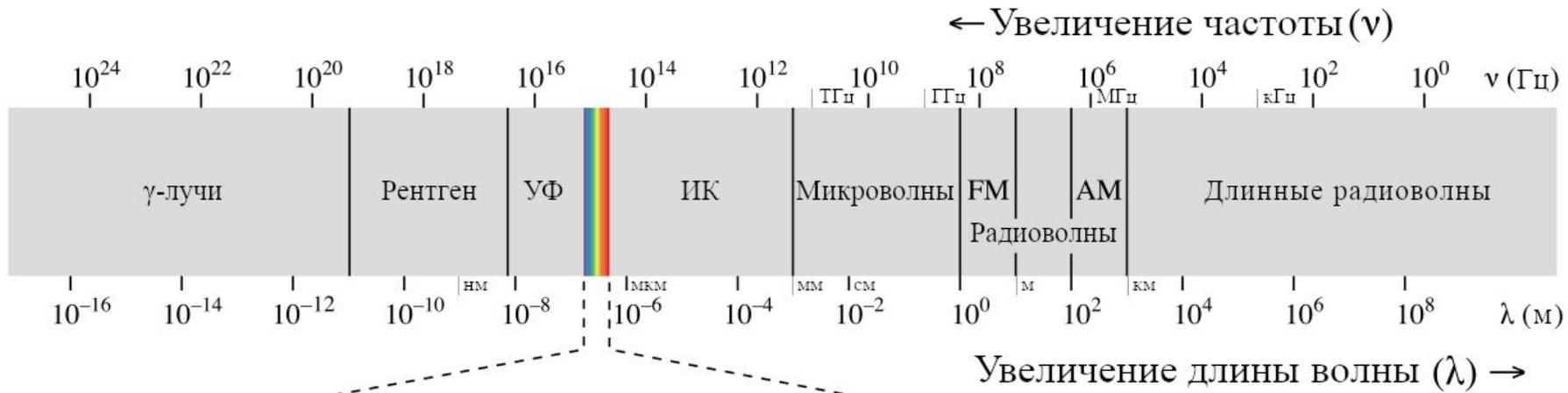


(j) Small Interclass Variations: four different categories

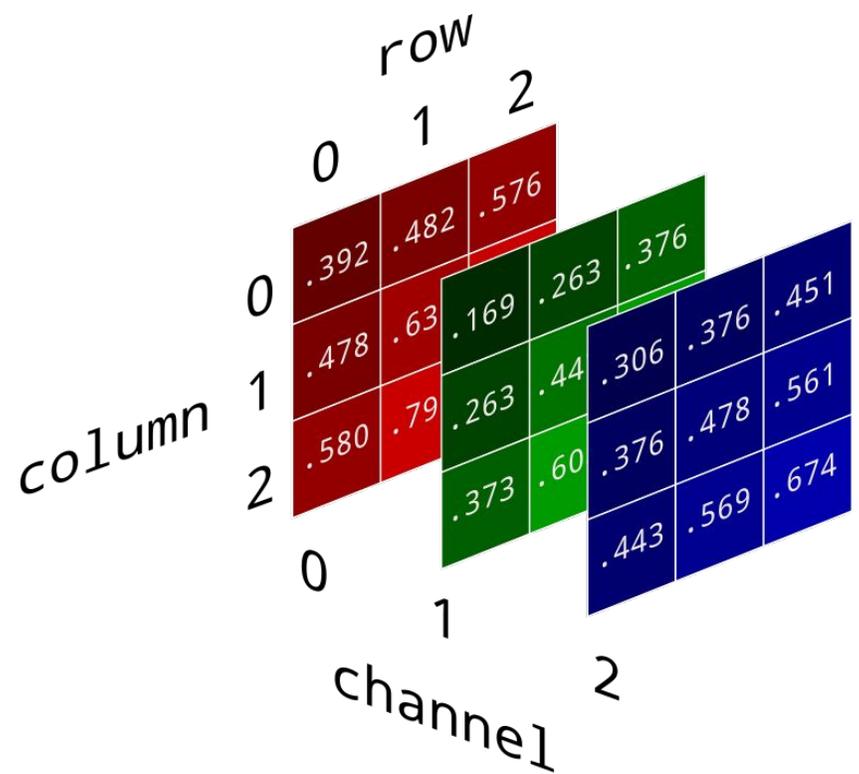
Способы представления изображений и видео



Физика



Представление картинок



Форматы хранения картинок



Exceptional.

Maximum quality JPEG, 300 dpi,
CMYK.



Poor!

Low-to-Medium quality JPEG, 150 dpi,
RGB.



JPEG



PNG



GIF

Введение в машинное обучение



Машинное обучение

CRISP-DM

Cross-industry standard process for data mining



Машинное обучение. Основные типы задач

Обучение с учителем (supervised learning)

- Каждый прецедент представляет собой пару «объект, ответ». Требуется найти функциональную зависимость ответов от описаний объектов и построить алгоритм, принимающий на входе описание объекта и выдающий на выходе ответ.

Обучение без учителя (unsupervised learning)

- Ответы не задаются, и требуется искать зависимости между объектами

Частичное обучение (semi-supervised learning)

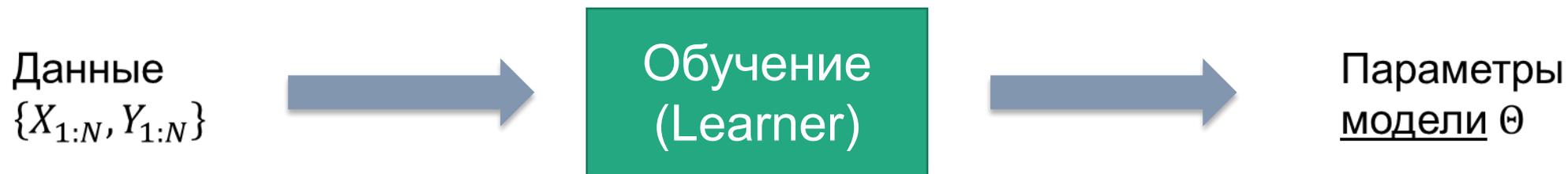
- Каждый прецедент представляет собой пару «объект, ответ», но ответы известны только на части прецедентов.

Обучение с подкреплением (reinforcement learning)

- Роль объектов играют пары «ситуация, принятое решение», ответами являются значения функционала качества, характеризующего правильность принятых решений (реакцию среды).

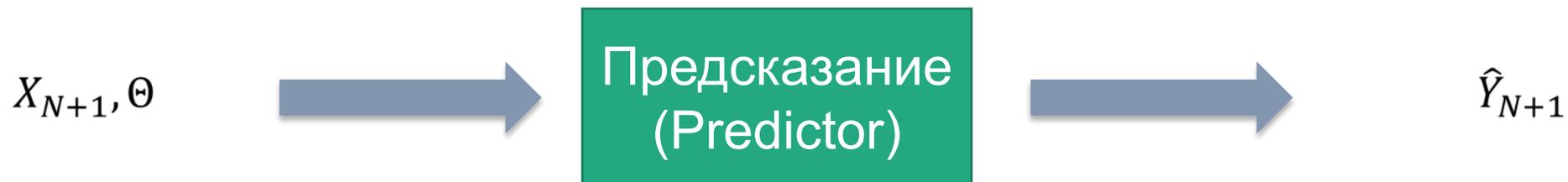
Машинное обучение. Обучение с учителем

Этап обучения (train)



Необходимо учитывать представительность выборки

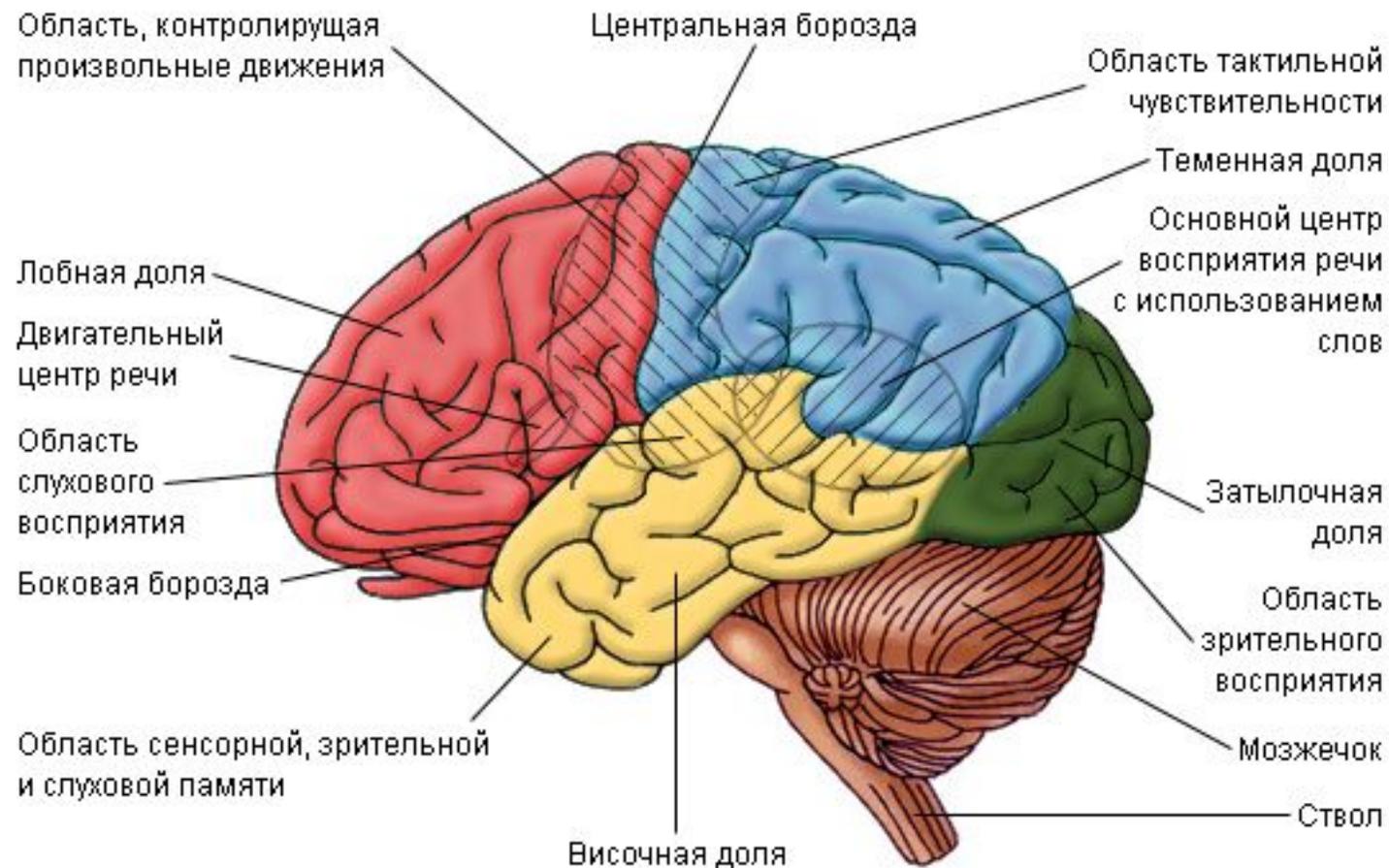
Этап применения (test)



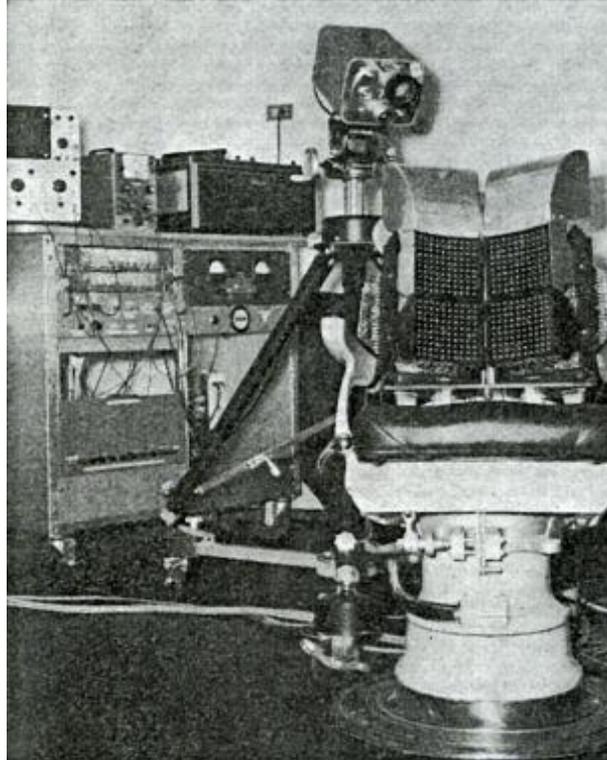
Нейронные сети в компьютерном зрении



Структура мозга человека

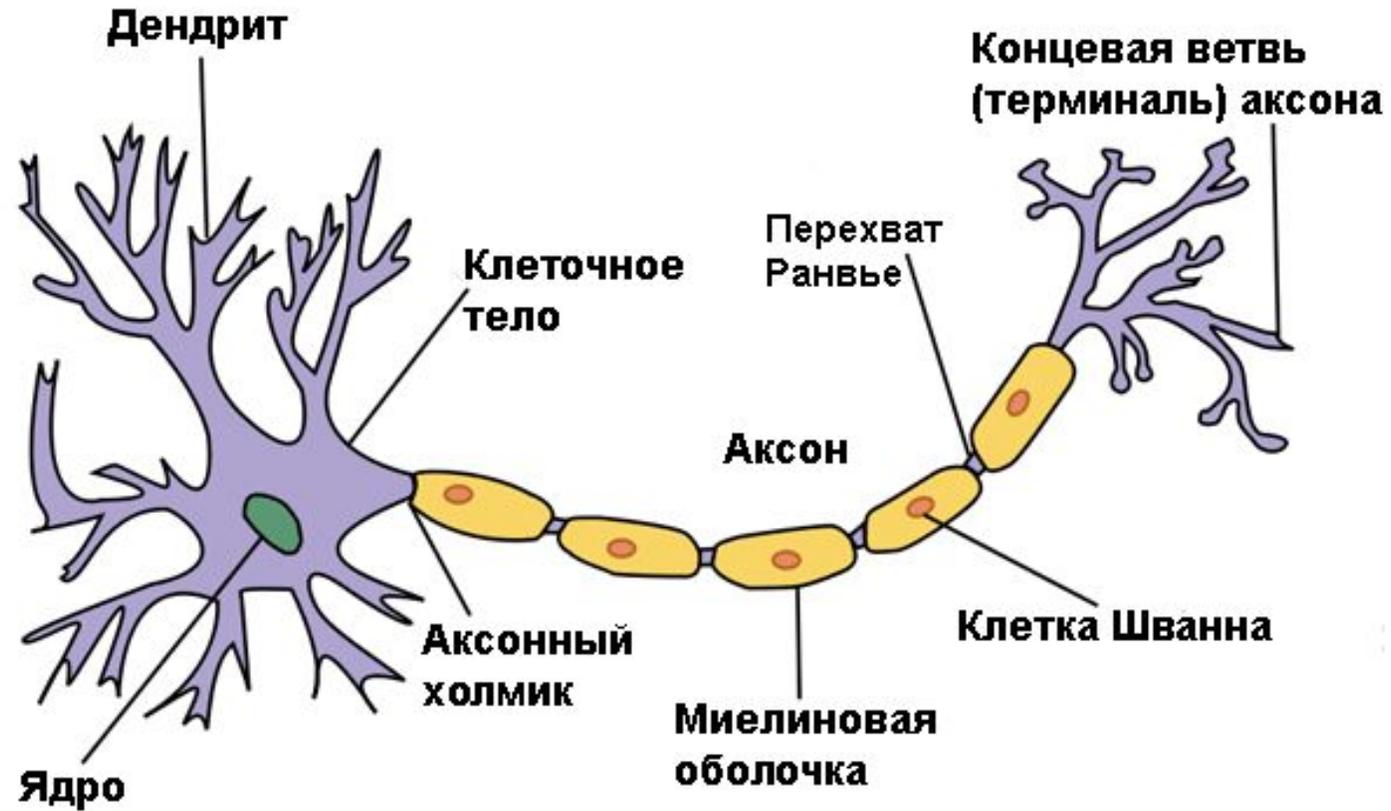


Нейропластичность

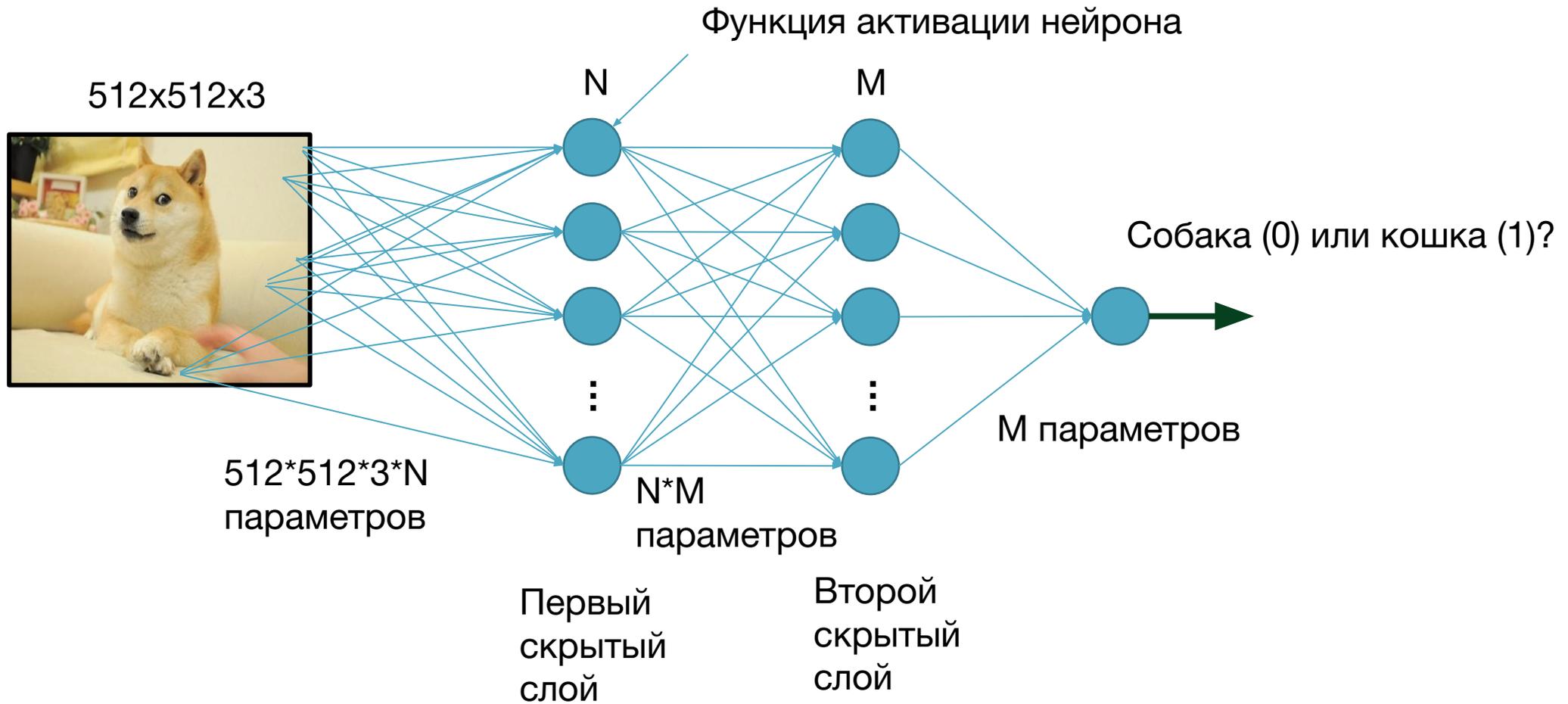


Пол Бак-и-Рита (Paul Bach-y-Rita)
Создан прибор (BrainPort), основой которого является матрица электродов
(до 20 на 20 штук), присоединяемых к языку (1969)

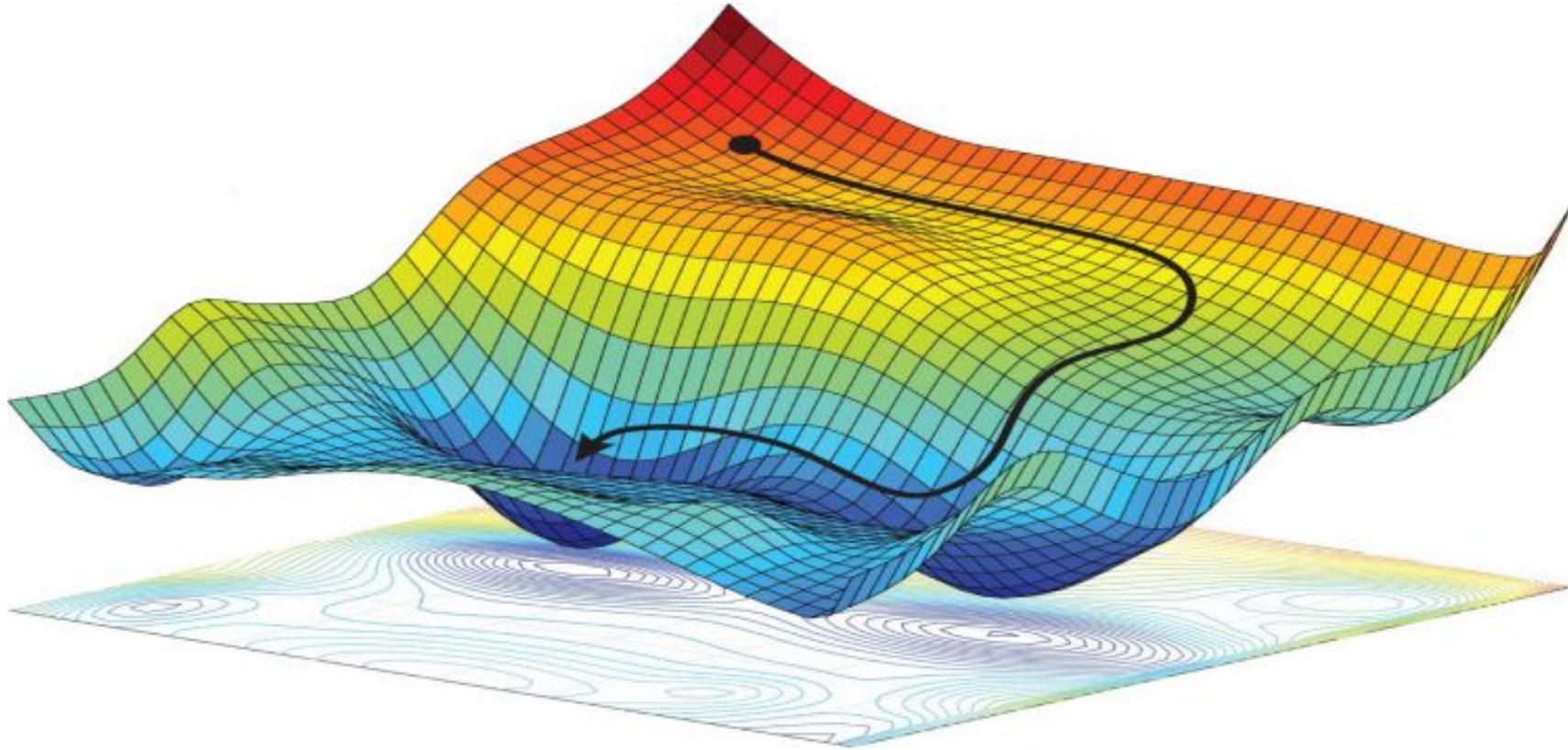
Типичная структура нейрона



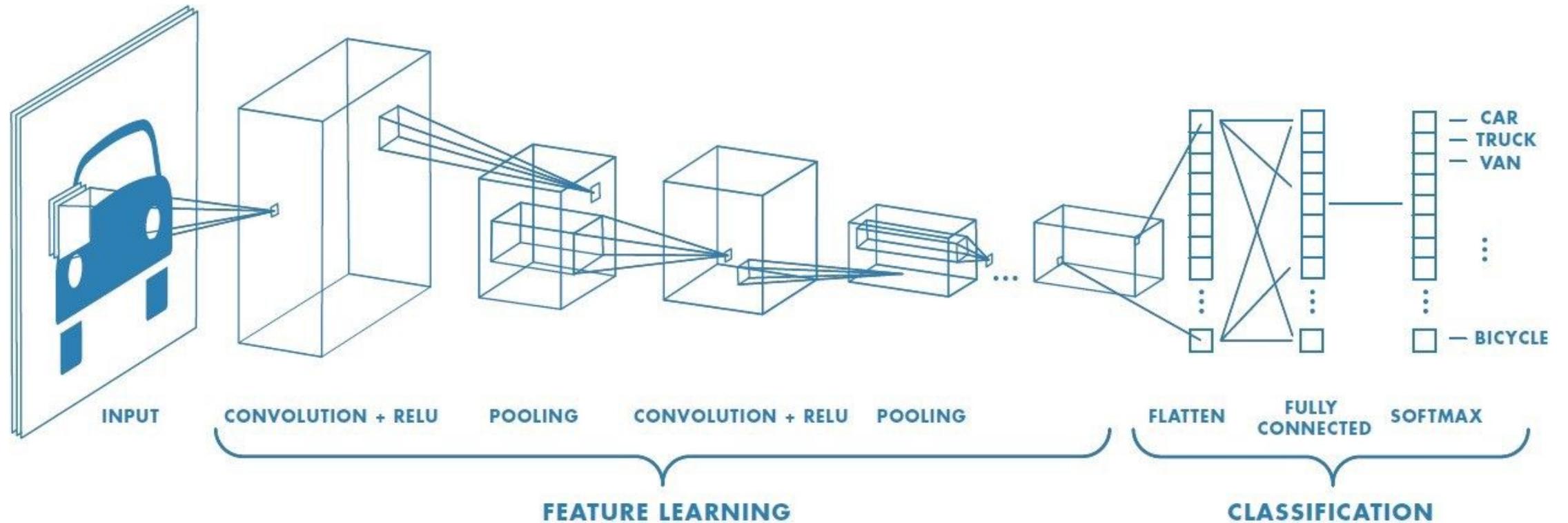
Многослойный персептрон



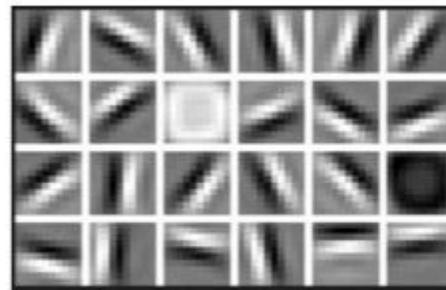
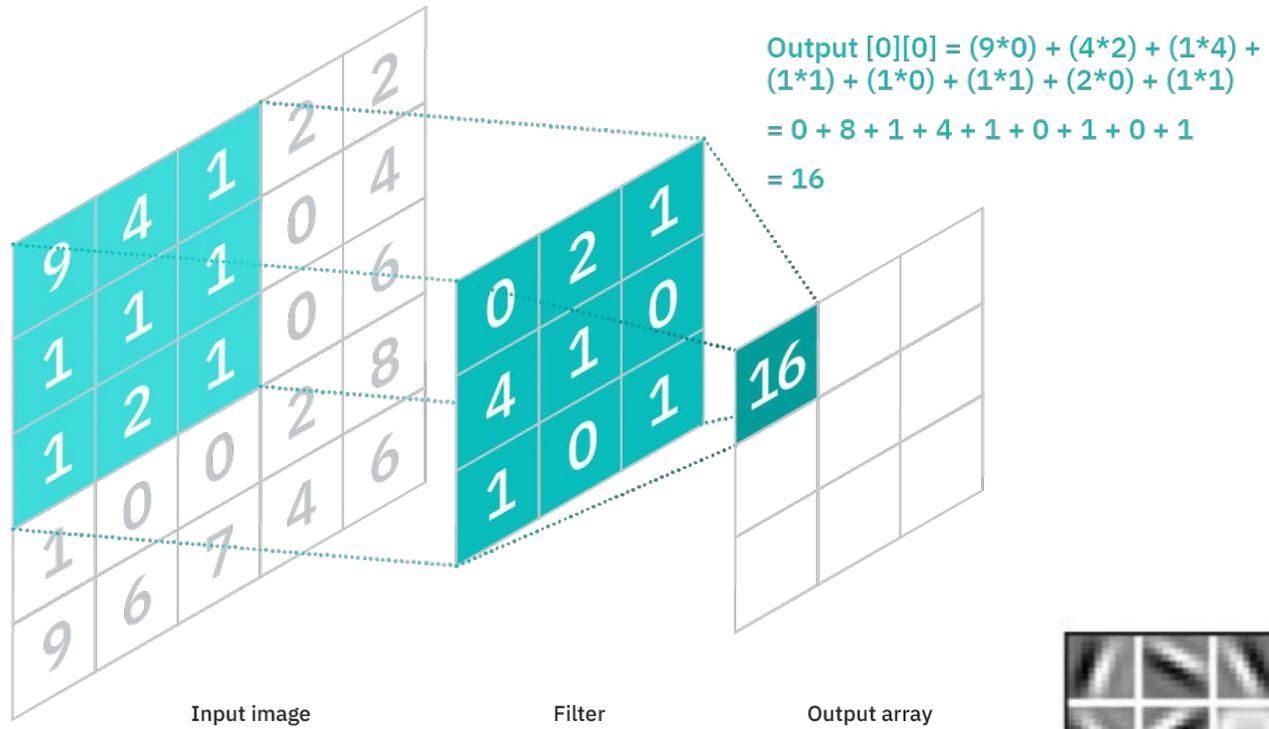
Оптимизация. Поиск параметров нейронной сети



Свёрточные нейронные сети (Convolutional Neural Network, CNN)



Свёртка



First Layer Representation

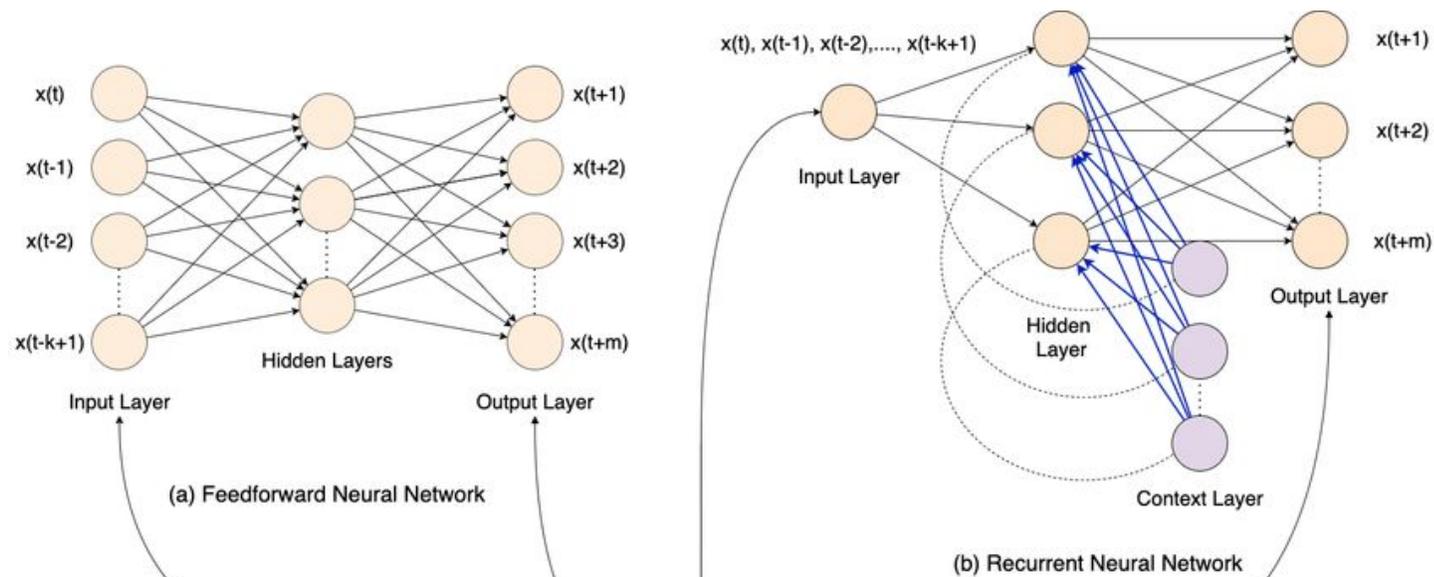


Second Layer Representation

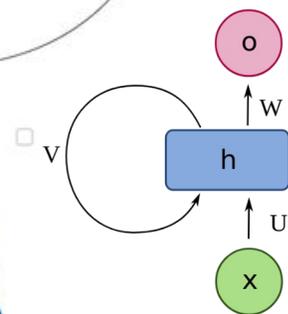
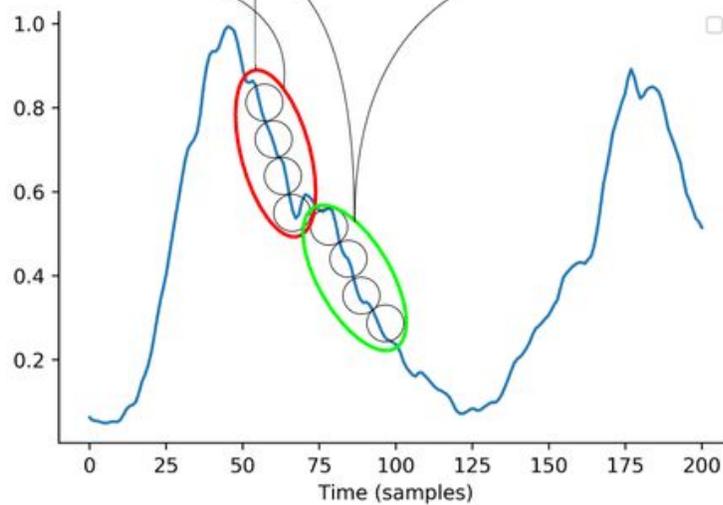


Third Layer Representation

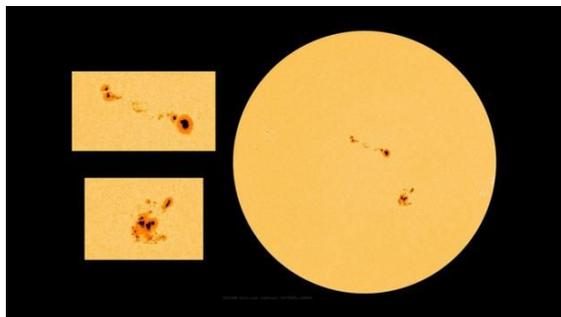
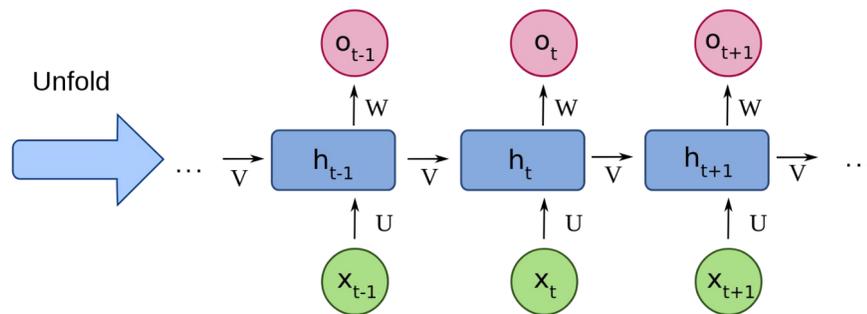
Рекуррентные нейронные сети



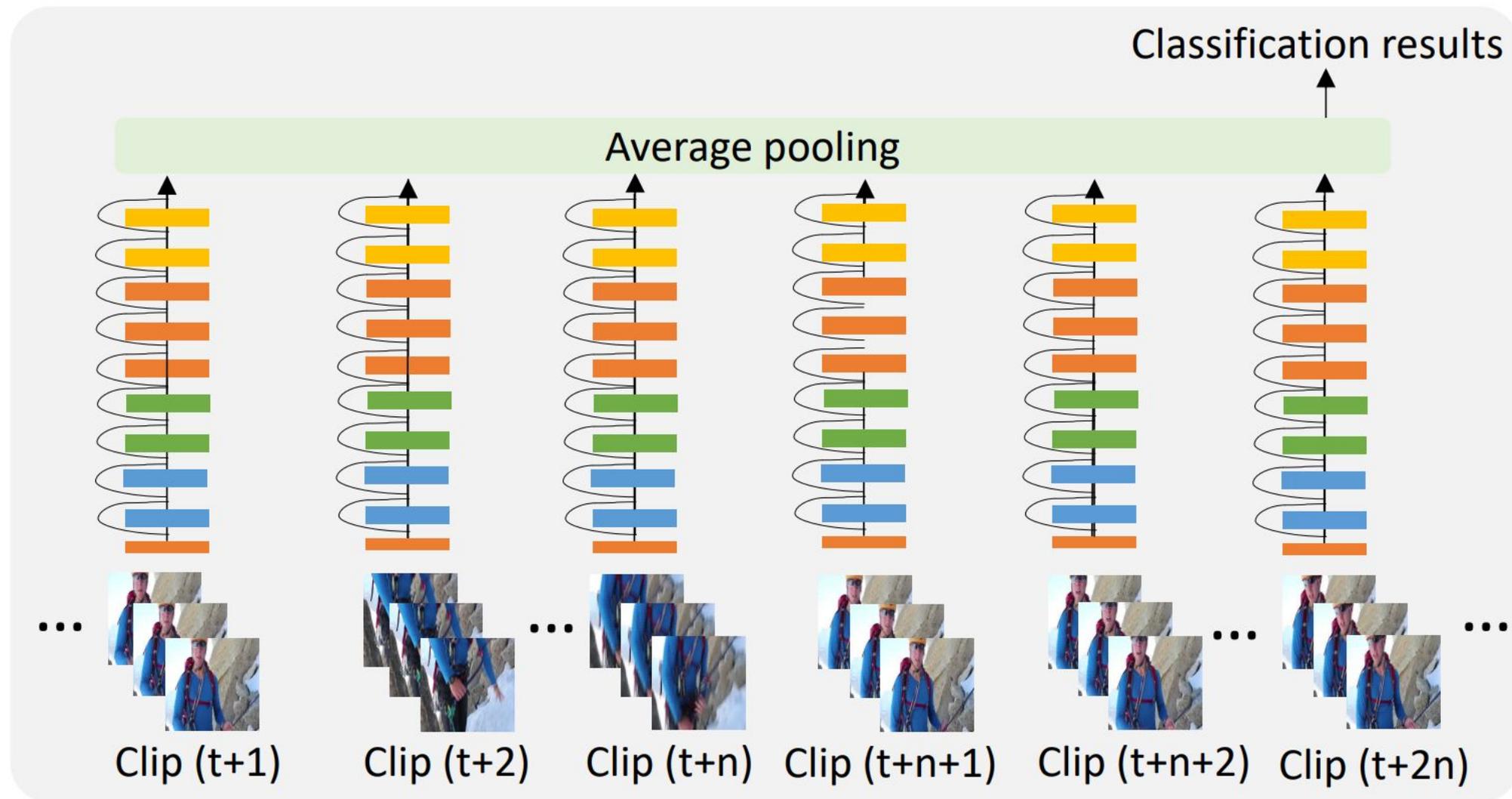
Sunspot Time-series



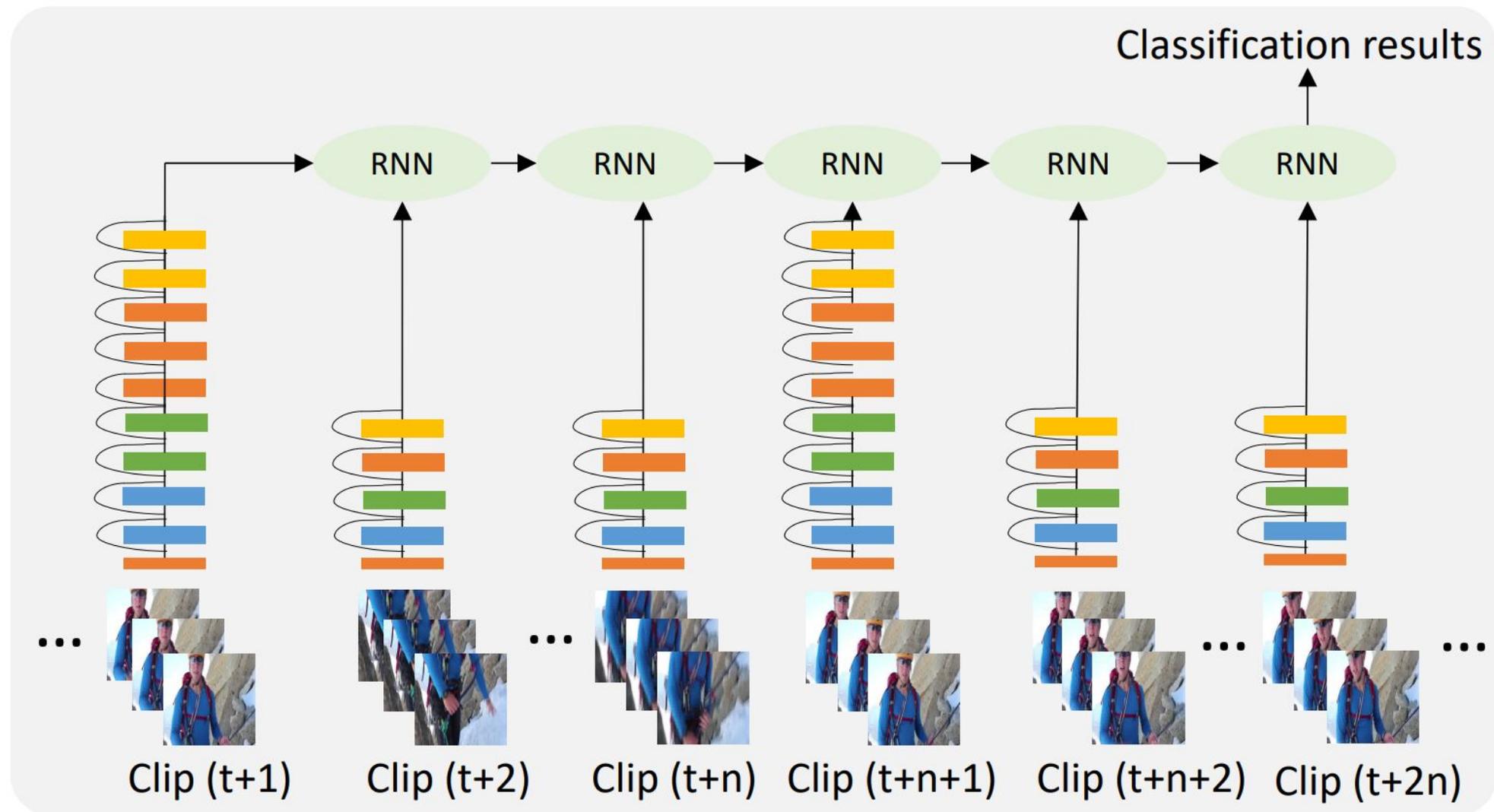
Unfold



Классификация видео. Пример 1



Классификация видео. Пример 2



Описание задания



Задача

- 1) Классификация видео (длительность 10 секунд)
- 2) Список категорий:
 - a) Спорт
 - b) Техника
 - c) Еда
 - d) Танцы
 - e) Животные



- 3) Инструкция по разметке данных:
<https://vk.cc/ciozJQ>



Спасибо
за внимание!