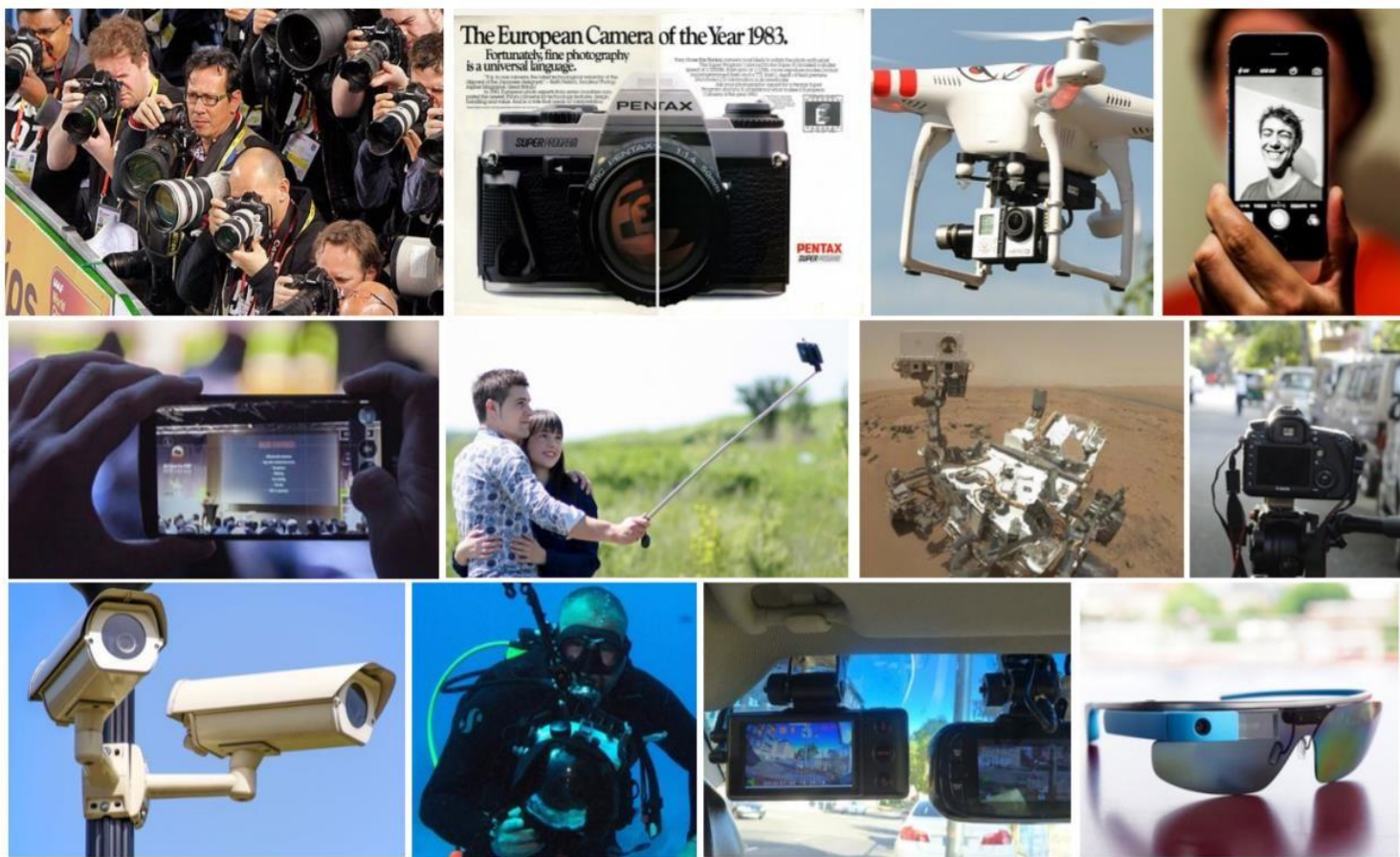


Введение в компьютерное зрение

Компьютерное зрение

Введение

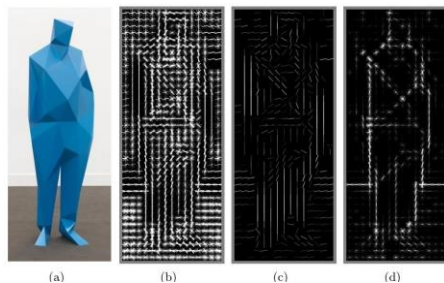


- 90% информации человек воспринимает при помощи зрения
 - 80% всего трафика в Интернете составляет видео
 - Что такое компьютерное зрение?
 - Что значит «воспринимать» изображение?
- «Воспринимать» ==
- Что и где?
 - Свойства?
 - Метрическая информация?

Компьютерное зрение

Определение

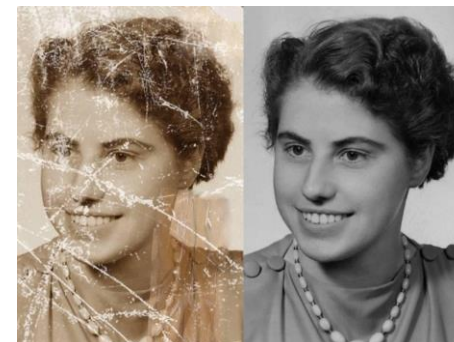
- Теория и технология создания машин, которые могут производить детекцию, отслеживание и классификация объектов
- Междисциплинарная область
- Обеспечение высокоуровневого понимания цифровых изображений и видео
- Автоматизация задач, которые может выполнять зрительная система человека
- Автоматическое извлечение, анализ и понимание полезной информации из изображения



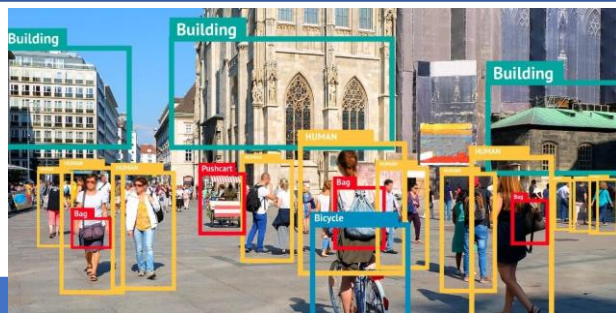
Компьютерное зрение

Задачи и методы

CV



Распознавание



Идентификация

Обнаружение

Классификация

Оценка
движения



Optical Flow

Отслеживание
объекта

Восстановление
изображений

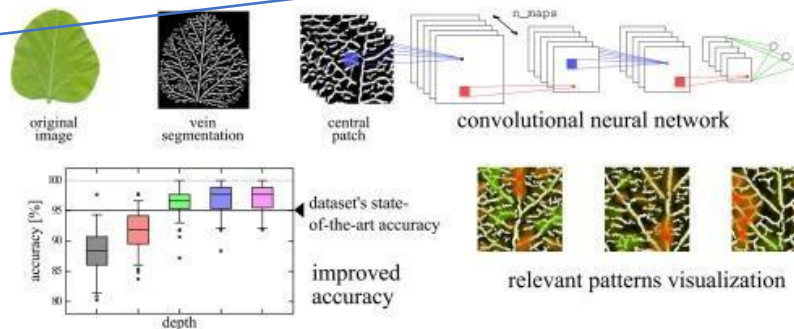
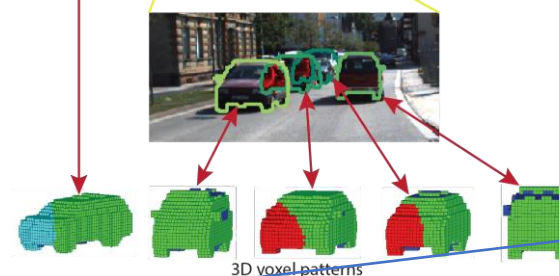
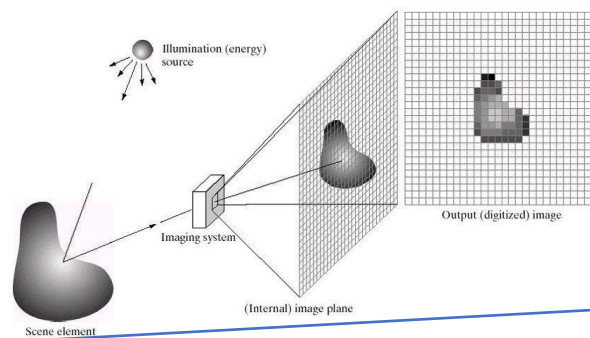
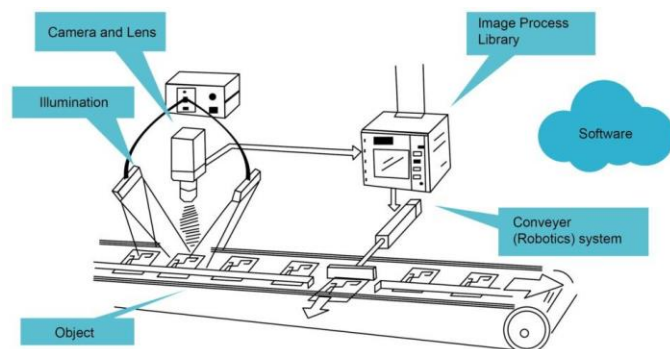
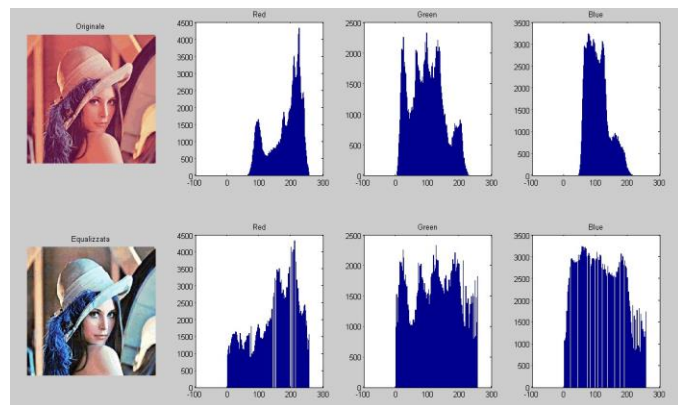
Реконструкция
изображений

Генерация
изображений

Восстановление
сцены

Компьютерное зрение

Разделы



Визуализация

Обработка изображений

Компьютерное зрение

Машинное зрение

Распознавание образов

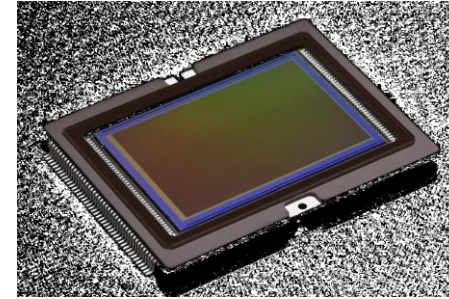
Компьютерное зрение

Связанные области

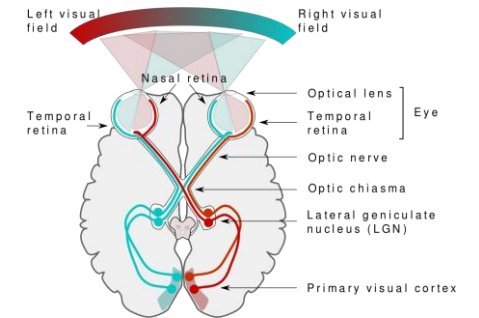


Искусственный
интеллект

Физика
твёрдого тела



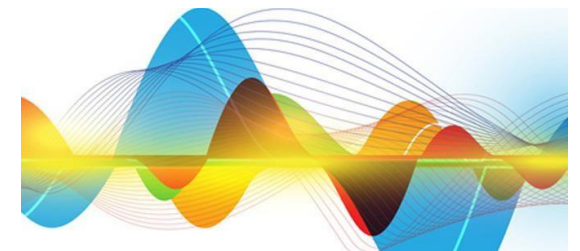
Нейробиология



Компьютерные
науки



Обработка
сигналов



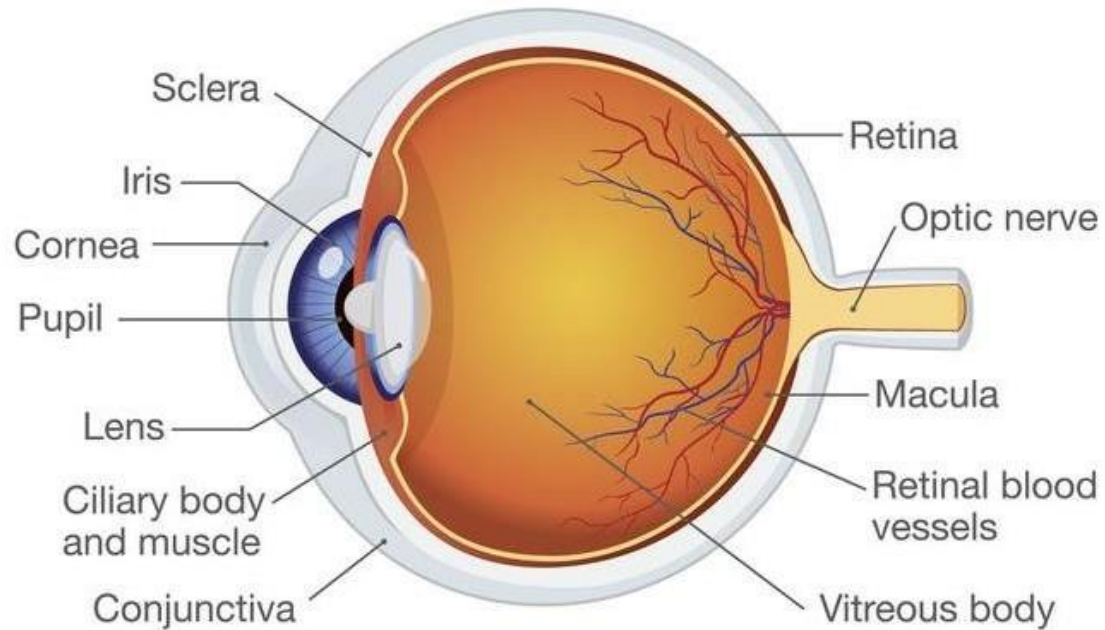
Компьютерное зрение

Основные этапы

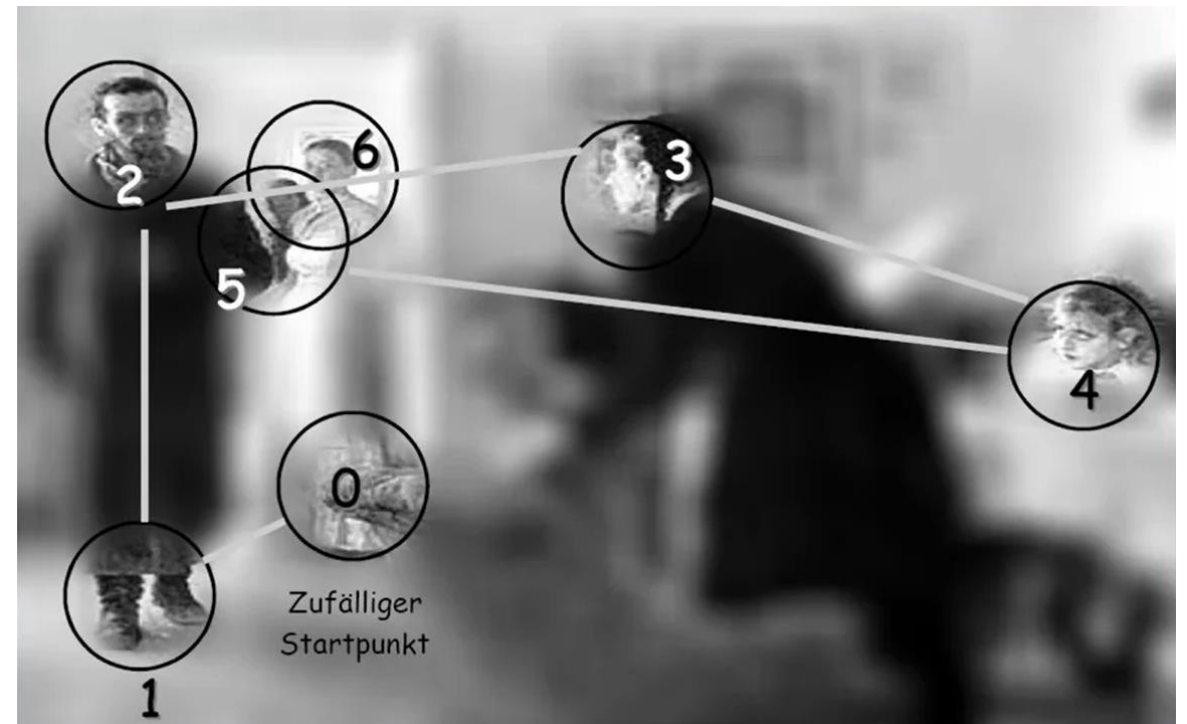


Компьютерное зрение

Биологические предпосылки



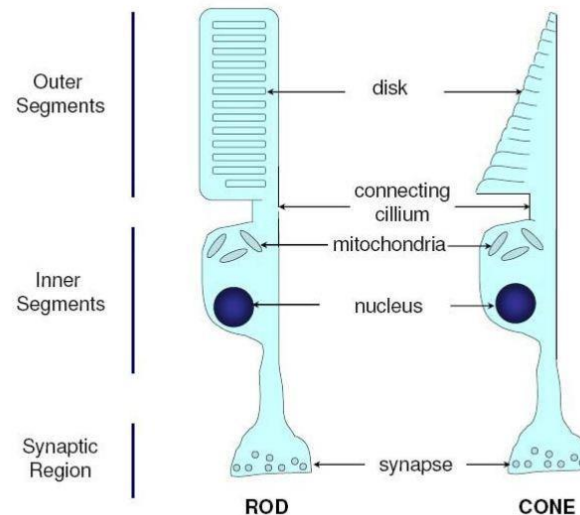
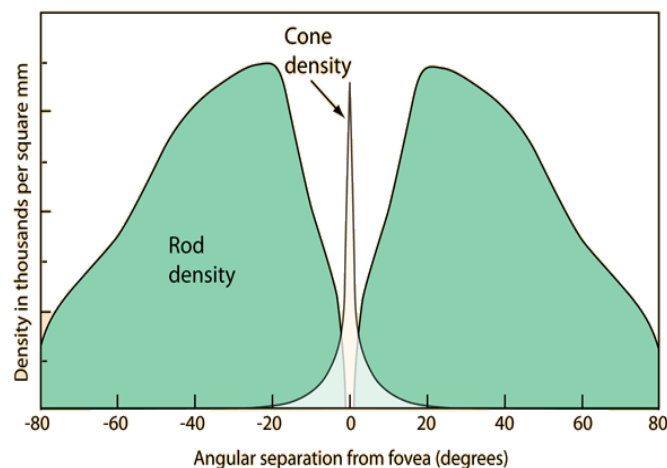
Устройство человеческого глаза



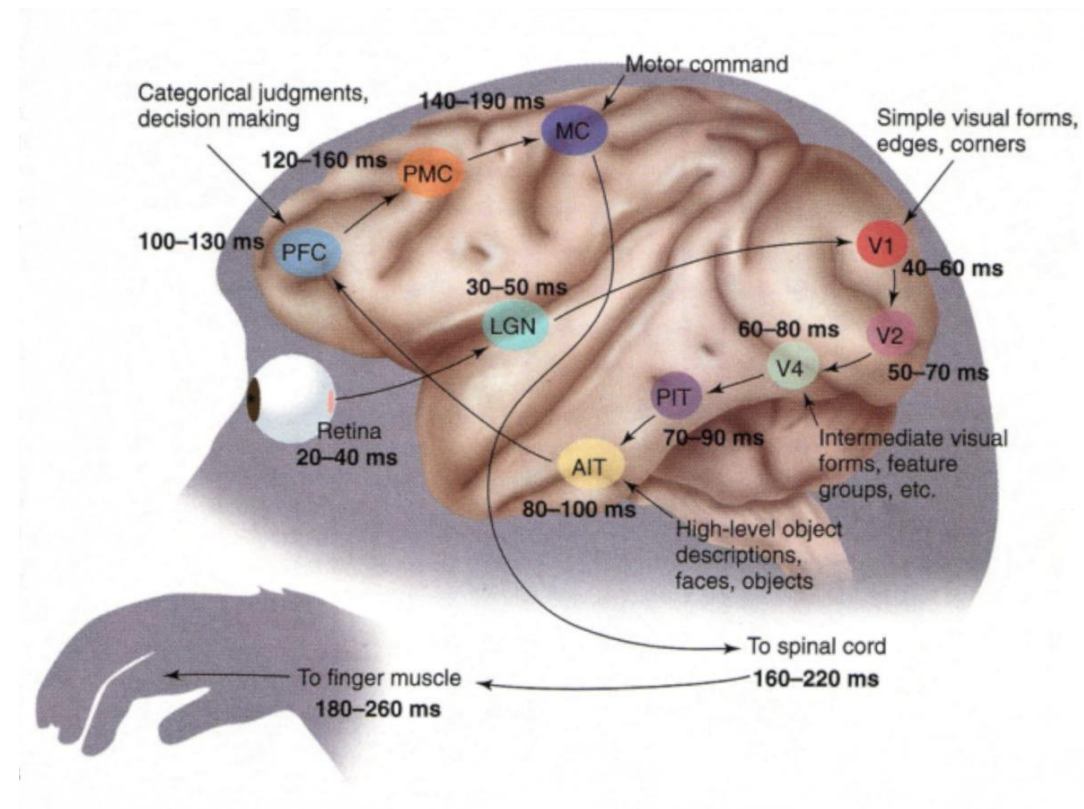
Исследование движения глаз (Yarbus A.L., 1967)

Компьютерное зрение

Биологические предпосылки



Неоднородность расположения палочек и колбочек
Палочки – яркость, колбочки – цвет



Распространение нервных импульсов

Компьютерное зрение

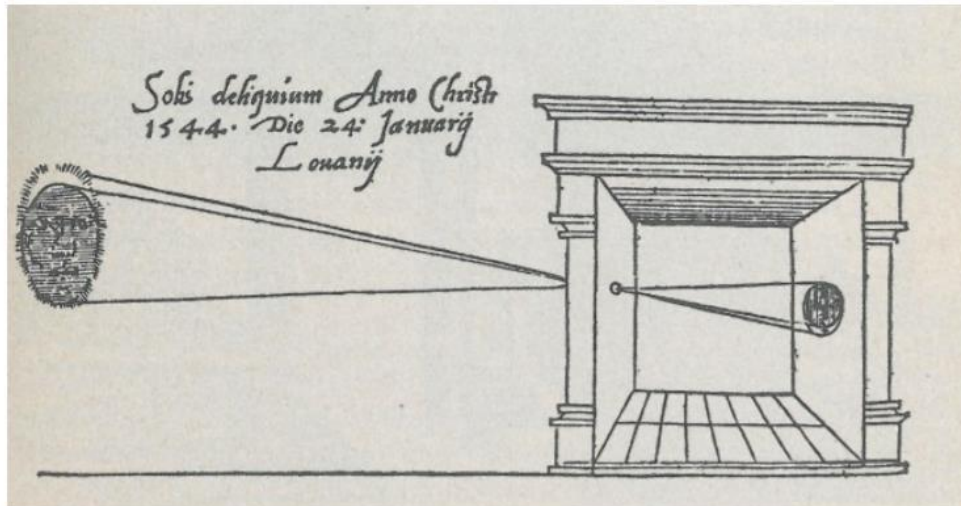
Биологические предпосылки



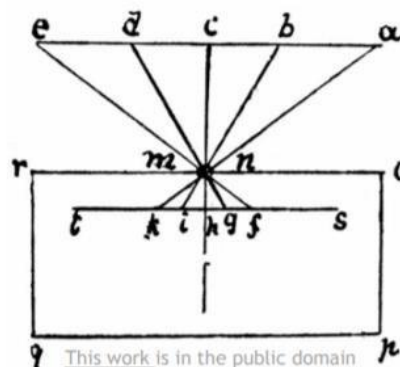
Компьютерное зрение

История развития

Gemma Frisius, 1545



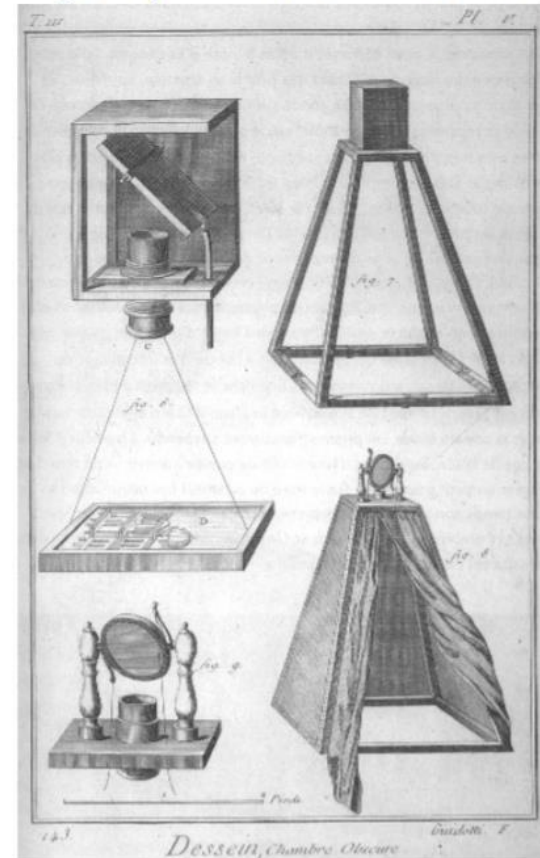
[This work is in the public domain](#)



Leonardo da Vinci,
16th Century AD

[This work is in the public domain](#)

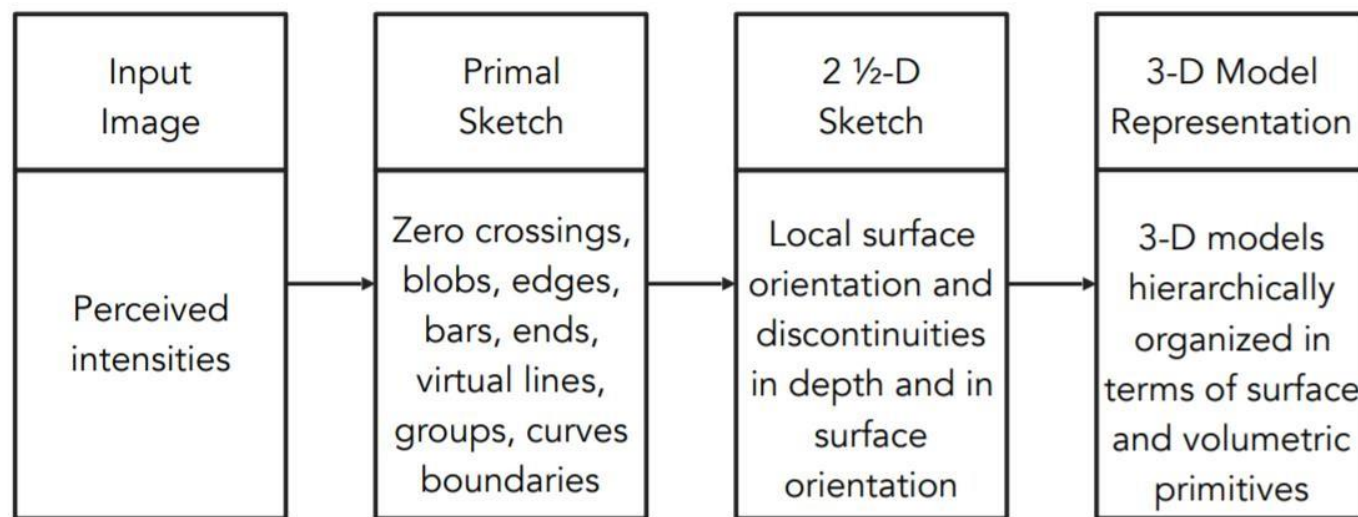
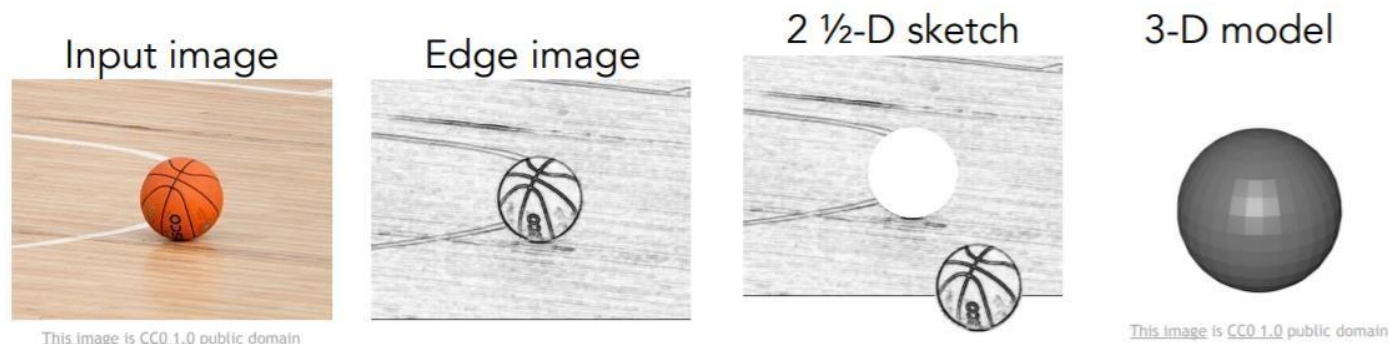
Encyclopedie, 18th Century



[This work is in the public domain](#)

Компьютерное зрение

История развития



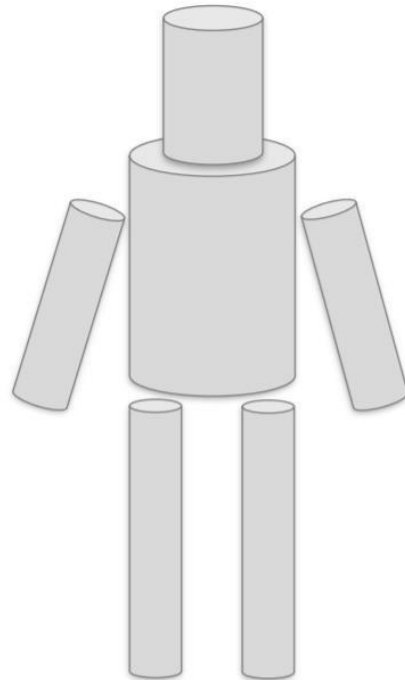
Stages of Visual Representation, David Marr, 1970s

Компьютерное зрение

История развития

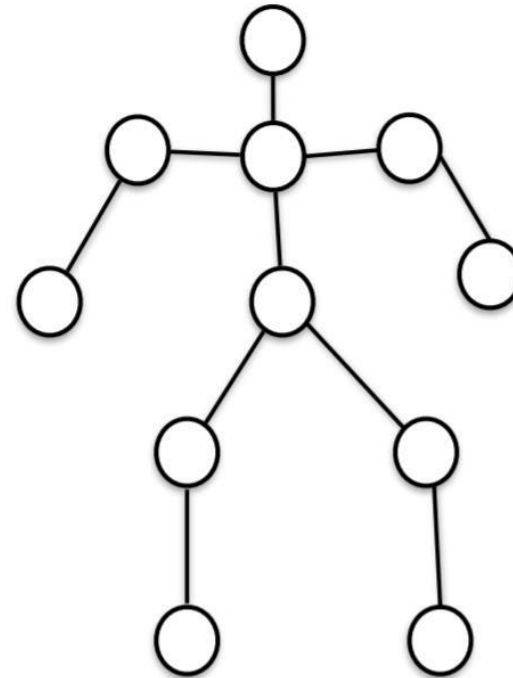
- Generalized Cylinder

Brooks & Binford, 1979



- Pictorial Structure

Fischler and Elschlager, 1973



Компьютерное зрение

История развития

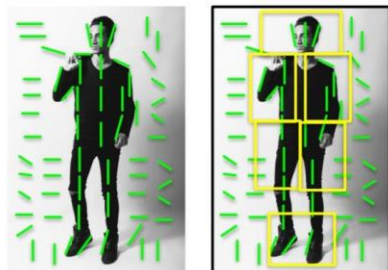


Image is public domain



Image is CC BY-SA 2.0

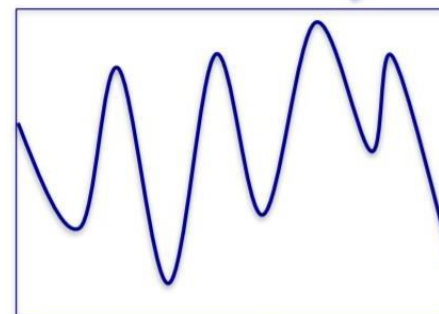
"SIFT" & Object Recognition, David Lowe, 1999



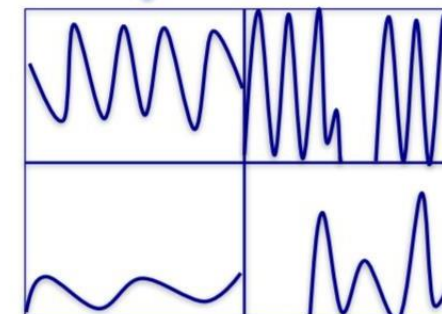
Deformable Part Model
Felzenswalb, McAllester, Ramanan, 2009



Image is CC0 1.0 public domain



Level 0



Level 1

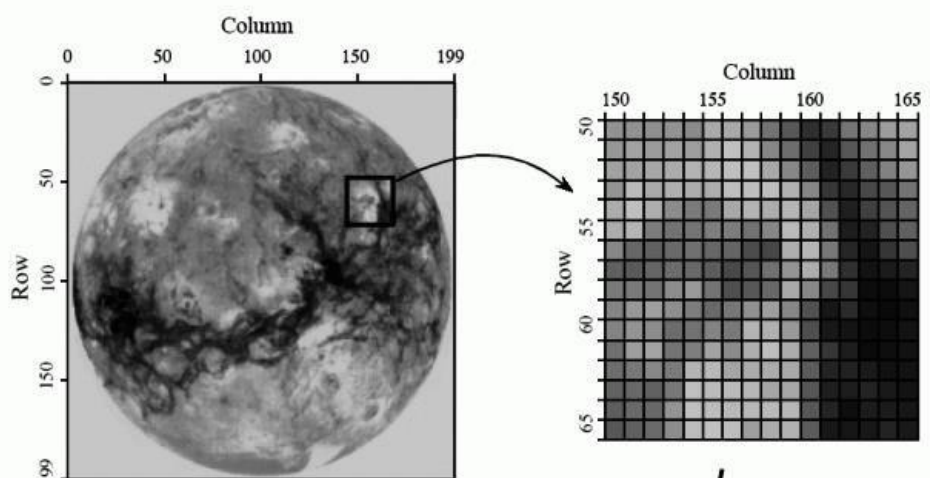
Spatial Pyramid Matching, Lazebnik, Schmid & Ponce, 2006

Цифровая обработка изображений

Введение

Изображение:

- Двумерная функция $f(x, y)$, x, y – координаты в пространстве, f – значение интенсивности пикселя
- Цифровое == x, y, f принимают конечное число дискретных значений



		Column															
		150	155	160	165												
Row	50	183	183	181	184	177	200	200	189	159	135	94	105	160	174	191	196
	51	186	195	190	195	191	205	216	206	174	153	112	80	134	157	174	196
55	52	194	196	198	201	206	209	215	216	199	175	140	77	106	142	170	186
	53	184	212	200	204	201	202	214	214	205	173	102	84	120	134	159	
60	54	202	215	203	179	165	165	199	207	202	208	197	129	73	112	131	146
	55	203	208	166	159	160	168	166	157	174	211	204	158	69	79	127	143
65	56	174	149	143	151	156	148	146	123	118	203	208	162	81	58	101	125
	57	143	137	147	153	150	140	121	133	157	184	203	164	94	56	66	80
60	58	164	165	159	179	188	159	126	134	150	199	174	119	100	41	41	58
	59	173	187	193	181	167	151	162	182	192	175	129	60	88	47	37	50
65	60	172	184	179	153	158	172	163	207	205	188	127	63	56	43	42	55
	61	156	191	196	159	167	195	178	203	214	201	143	101	69	38	44	52
65	62	154	163	175	165	207	211	197	201	201	199	138	79	76	67	51	53
	63	144	150	143	162	215	212	211	209	197	198	133	71	69	77	63	53
65	64	140	151	150	185	215	214	210	210	211	209	135	80	45	69	66	60
	65	135	143	151	179	213	216	214	191	201	205	138	61	59	61	77	63

Уровни процессов:

- Низкий уровень – примитивные операции предобработки
- Средний уровень – сегментация, выделение признаков, сжатие информации
- Высокий уровень – осмысление набора распознанных объектов, принятие решений, интерпретация

Цифровая обработка изображений

История развития



Цифровое изображение, полученное в 1921 г. с кодовой ленты на телеграфном аппарате с особым шрифтом



Цифровое изображение, полученное в 1922 г. с использованием перфоленды после прохождения сигнала через Атлантику дважды



Неретушированная фотография генералов Першинга и Фоша, переданная в 1929 г. по кабелю из Лондона в Нью-Йорк с помощью 15-градационного оборудования

Цифровая обработка изображений

История развития

Развитие цифровой вычислительной техники

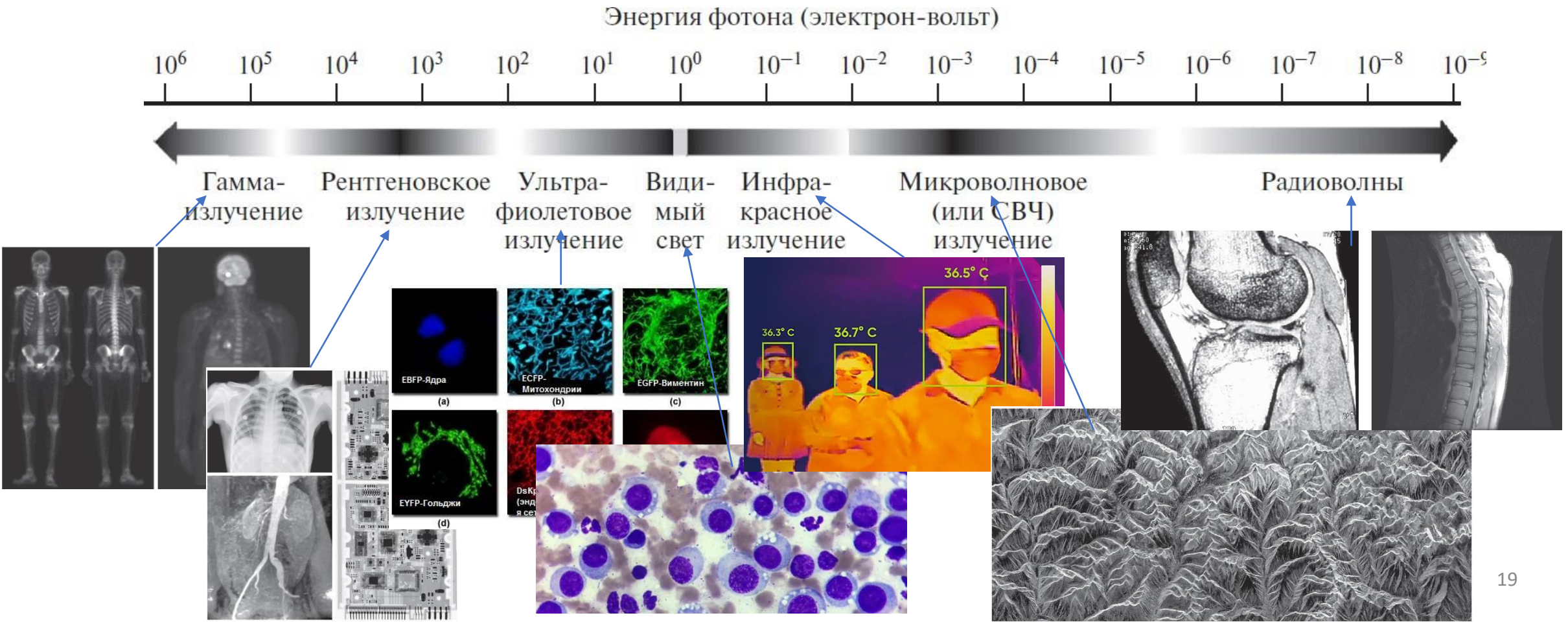


Джон фон Нейман



Цифровая обработка изображений

Формирование изображений



Цифровая обработка изображений

Основные этапы обработки



Распознавание образов

Оценка подходов

PASCAL Visual Object Challenge (20 object categories)

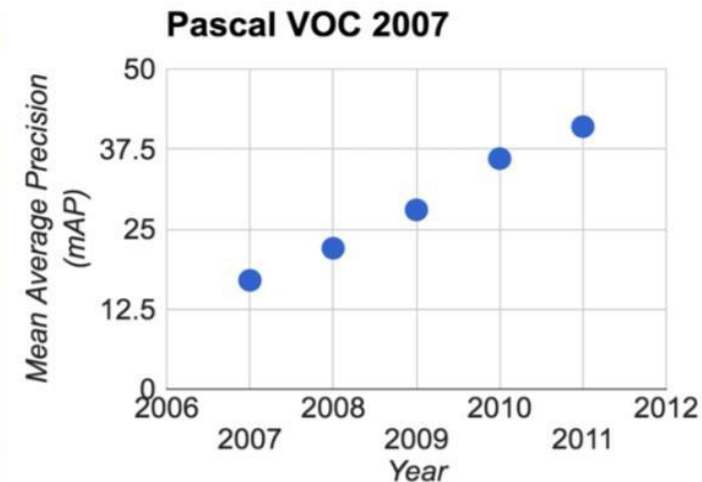
[Everingham et al. 2006-2012]



Image is CC0 1.0 public domain

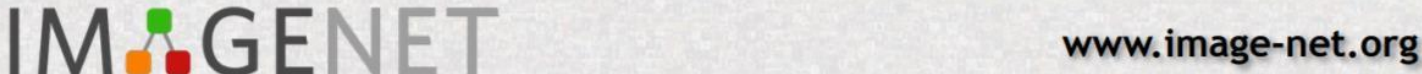



This image is licensed under
CC BY-SA 2.0; changes made



Распознавание образов

Оценка подходов



IMGENET www.image-net.org

22K categories and **14M** images

- Animals
 - Bird
 - Fish
 - Mammal
 - Invertebrate
- Plants
 - Tree
 - Flower
 - Food
 - Materials
- Structures
 - Artifact
 - Tools
 - Appliances
 - Structures
- Person
 - Scenes
 - Indoor
 - Geological Formations
 - Sport Activities

Deng, Dong, Socher, Li, Li, & Fei-Fei, 2009


Распознавание образов

Оценка подходов

IMAGENET Large Scale Visual Recognition Challenge

Steel drum

The Image Classification Challenge:
1,000 object classes
1,431,167 images



Output:
Scale
T-shirt
Steel drum
Drumstick
Mud turtle

✓

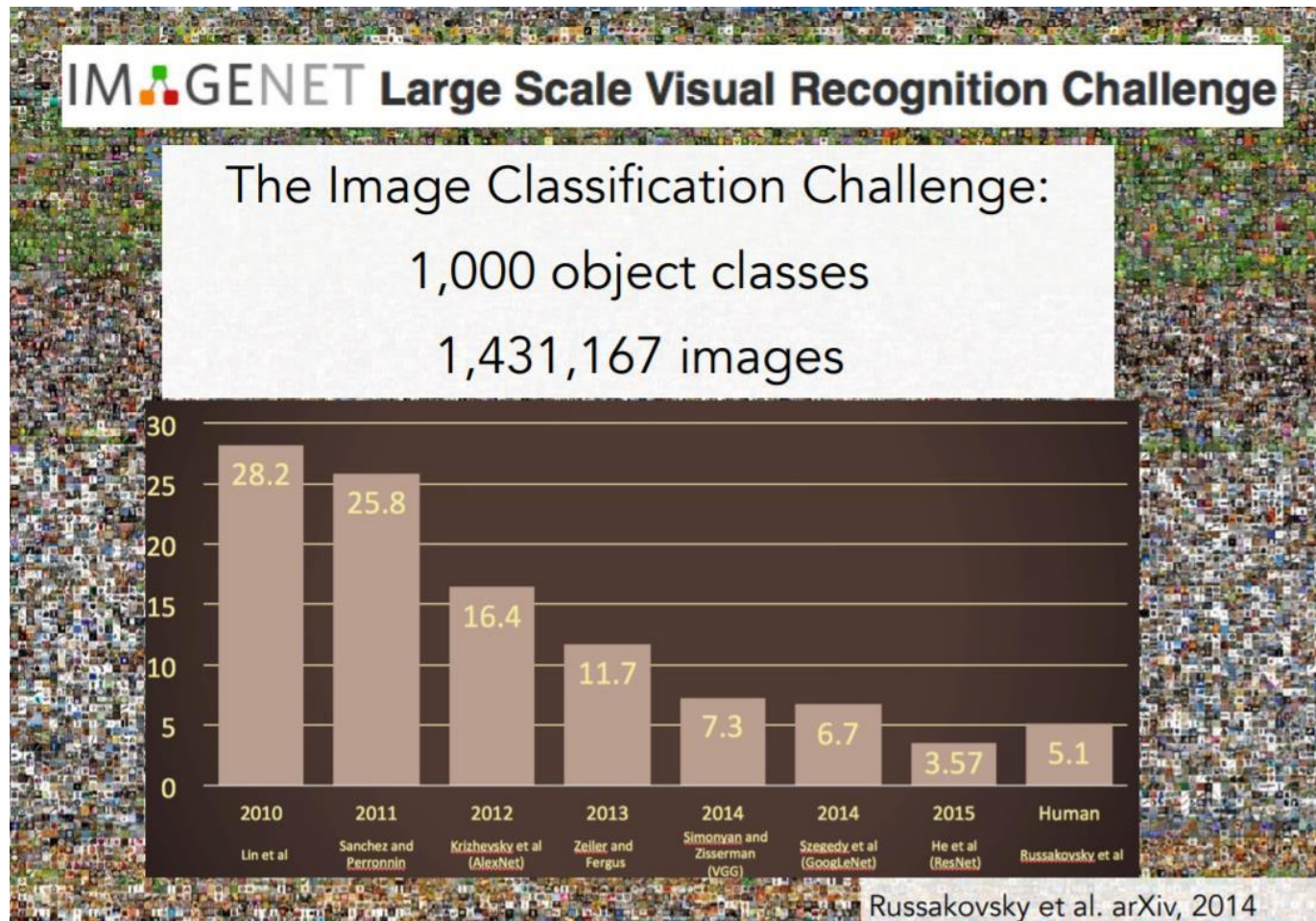
Output:
Scale
T-shirt
Giant panda
Drumstick
Mud turtle

✗

Russakovsky et al. arXiv, 2014

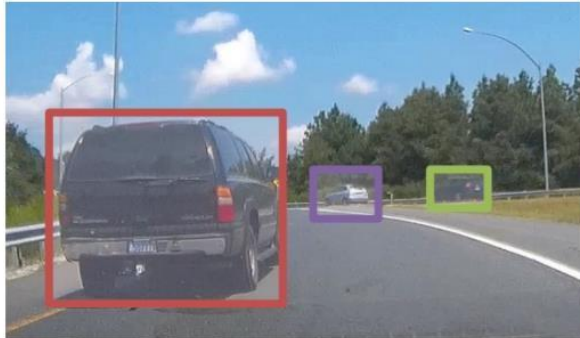
Распознавание образов

Оценка подходов



Распознавание образов

Современные задачи



This image is licensed under CC BY-NC-SA 2.0; changes made

- Object detection
- Action classification
- Image captioning
- ...

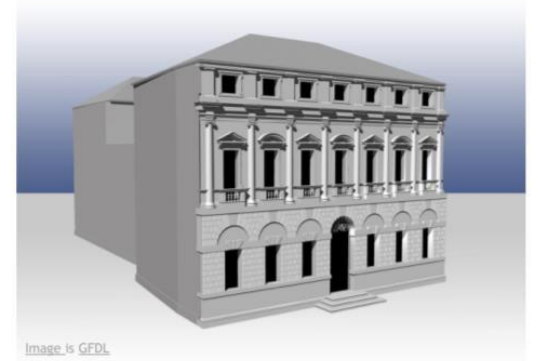
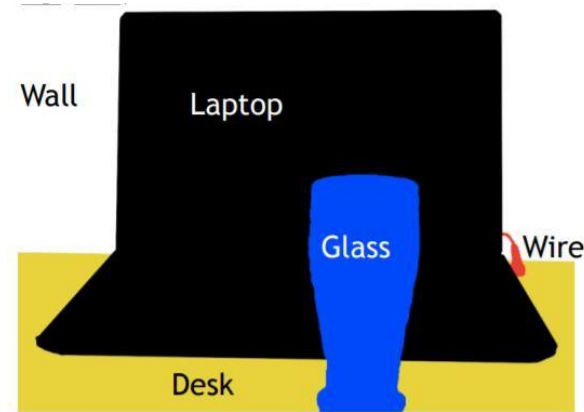
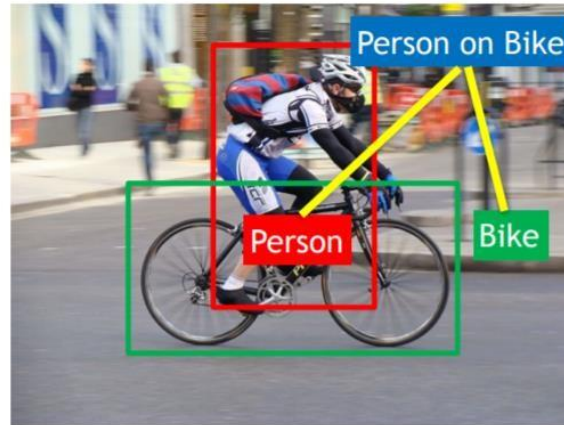


Image is GFDL



This image is licensed under CC BY-SA 2.0; changes made



This image is licensed under CC BY-SA 3.0; changes made

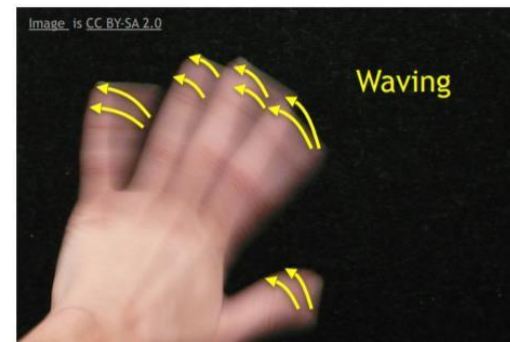


Image is CC BY-SA 2.0



Image is CC BY-SA 4.0

Распознавание образов

Сверточные нейронные сети

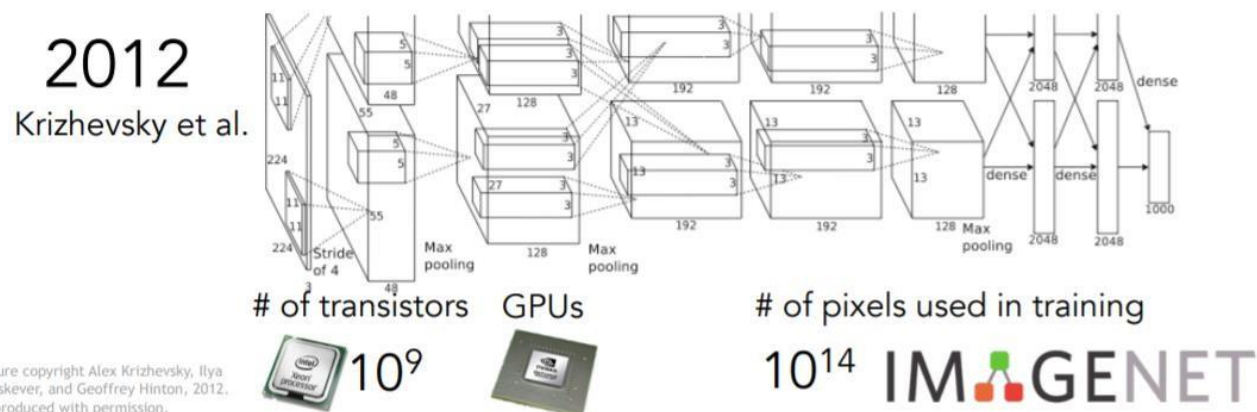
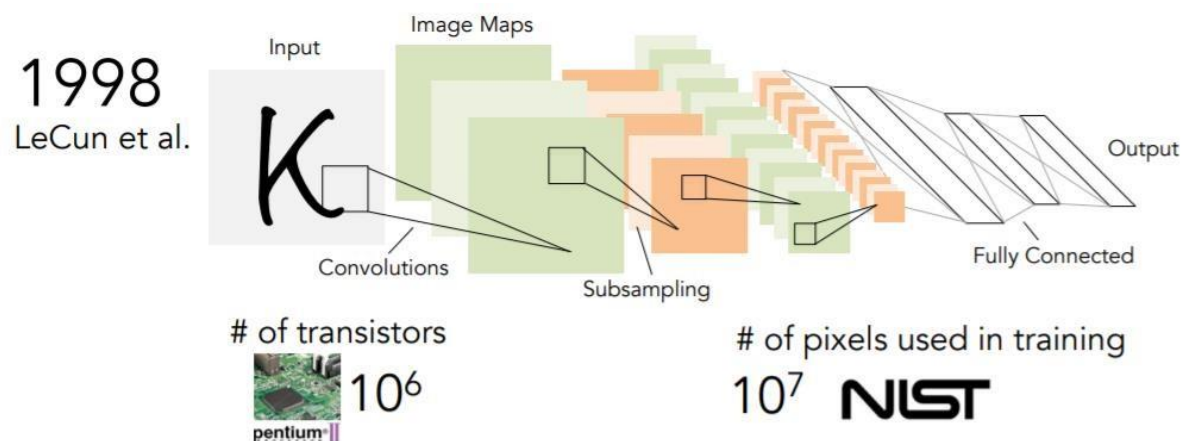


Figure copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

Резюме

- Обзор области компьютерного зрения
- Биологические предпосылки компьютерного зрения
- История развития компьютерного зрения
- Введение в цифровую обработку изображений
- Обзор области распознавания образов

Использованные материалы

- [Курс МФТИ по компьютерному зрению](#), автор – Колокольников Г.А.
- Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2012. – 1104 с. – ISBN 978-5- 94836-331-8.2.
- Курс лекций cs231n «Convolutional Neural Networks for Visual Recognition» (<http://cs231n.stanford.edu>).