

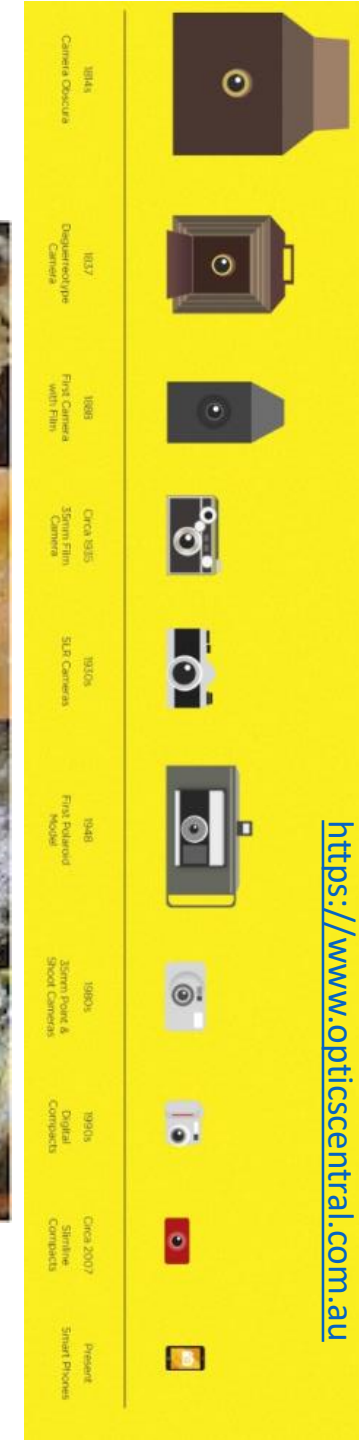
Эволюция зрения



(credit: Johannes Burge)

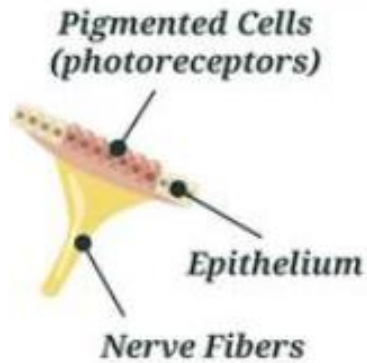
Александр Бовырин

<https://www.opticscentral.com.au>

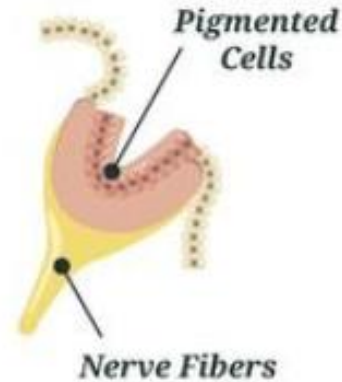


Эволюция глаза

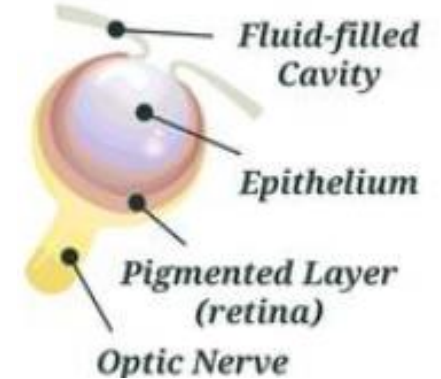
1. Светло или темно?



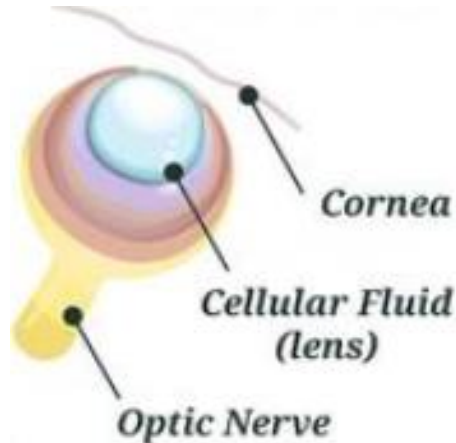
2. Где светло?



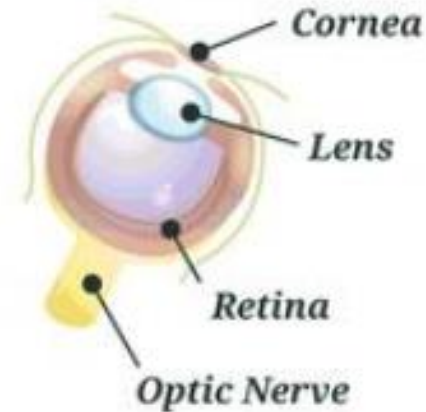
3. Камерный глаз.
Появляются расплывчатые образы.



4. Камерный глаз.
Появляются расплывчатые образы.



5. С линзой изображение становится резким.



Биноккулярное зрение и зрение в темноте



Почему хищники в основном обладают биноккулярным зрением?

Почему нехищникам выгоднее панорамное зрение?

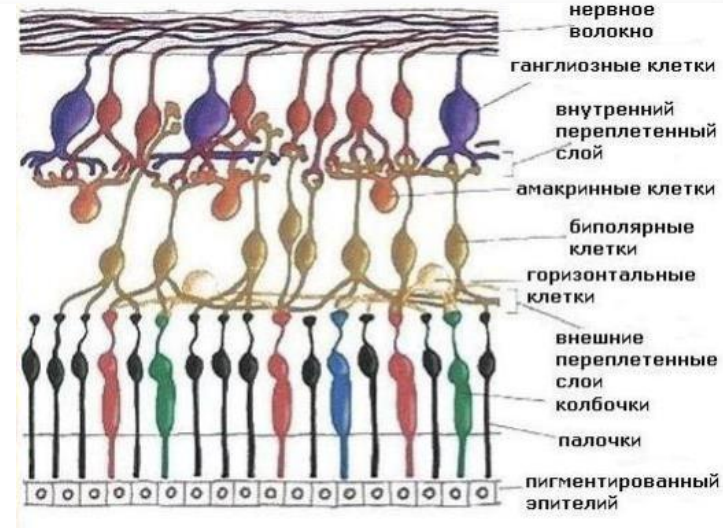


Почему глаза некоторых видов светятся в темноте?

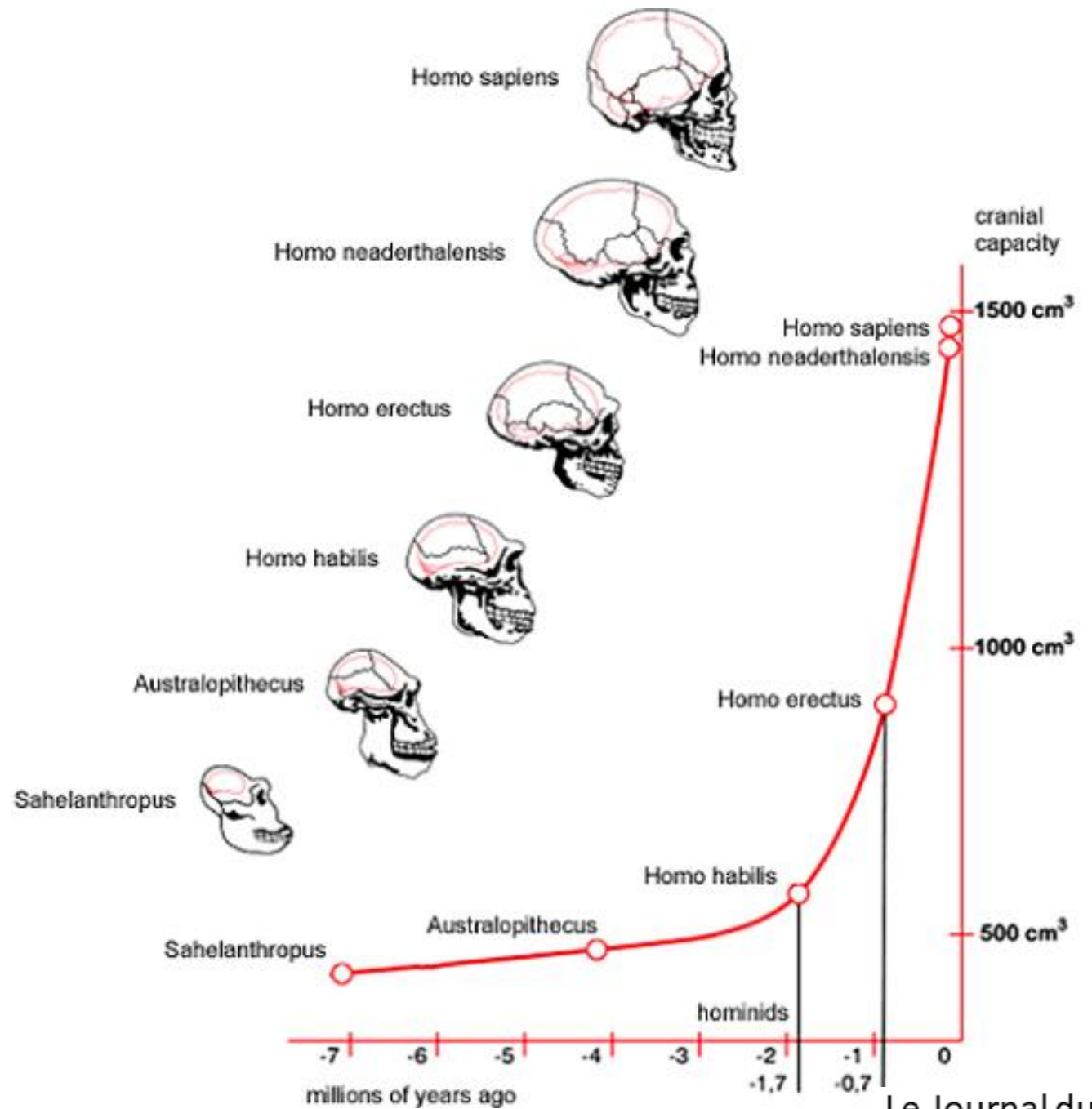
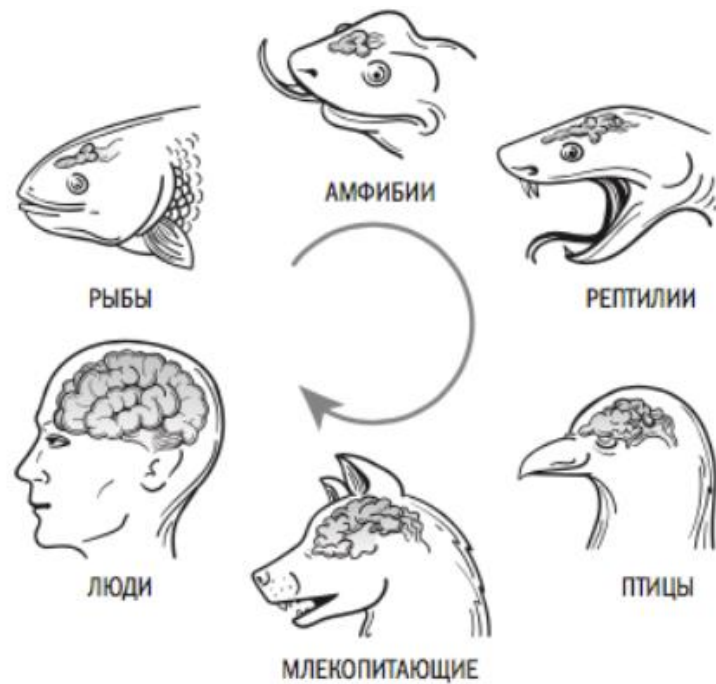


Глаз человека

- Зрачок 1..8 мм
- 100 млн палочек и 5 млн колбочек
- Видимый свет 330 нм -730 нм
- Распределение палочек не однородно. (средняя плотность $1,8 \cdot 10^5$ на 1 мм).
- Есть слепое пятно, где зрительный нерв
- Палочки могут реагировать на один фотон, но могут быть связаны только одним нейроном.
- У человека и приматов обнаружены колбочки с тремя разными кривыми спектральной чувствительности, максимумы которых у человека находятся в синей, зелёной и красной областях спектра.
- FPS (frames per second): лягушка 15—20 гц, человек — до 50—60 гц, муха — до 250—300 гц.

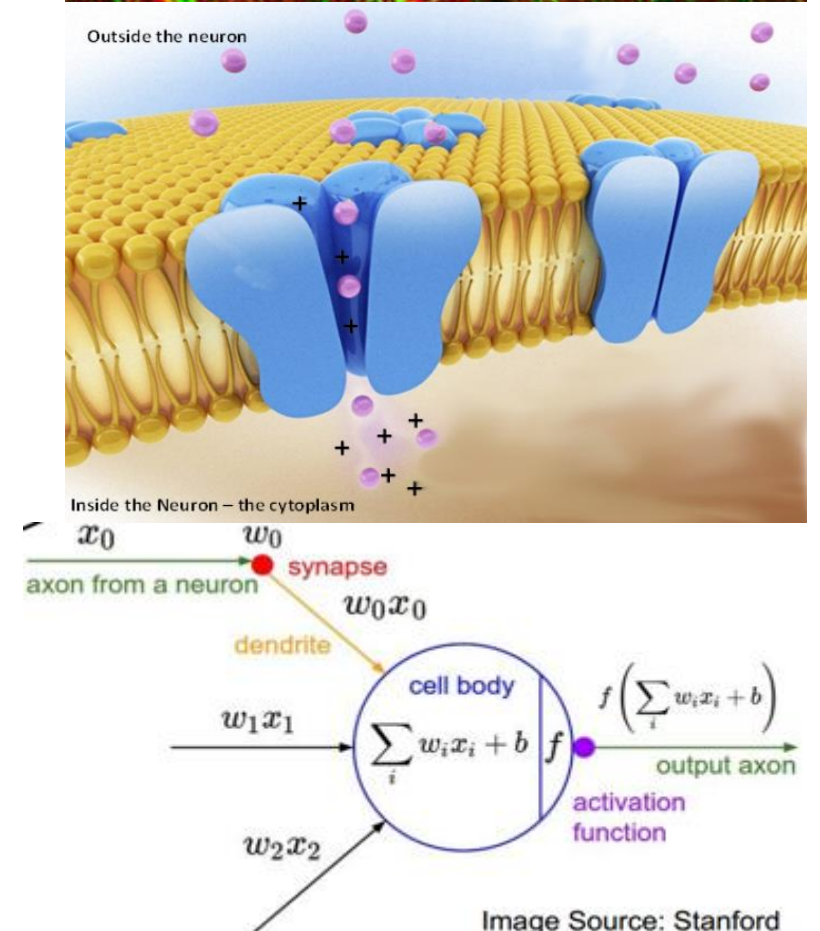
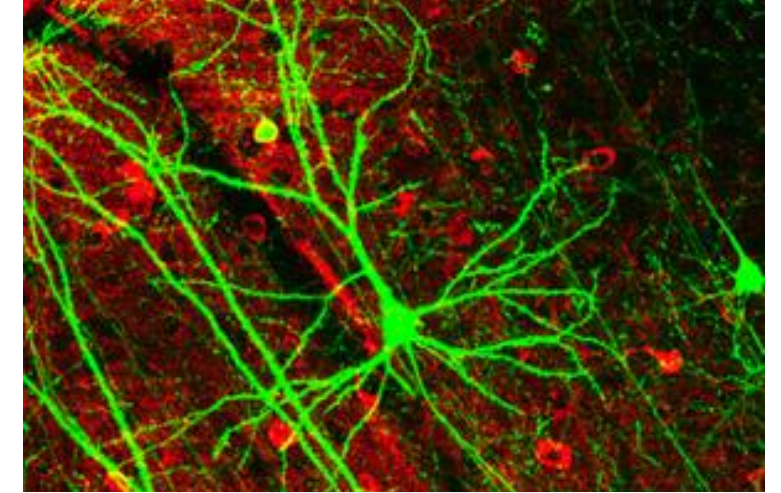
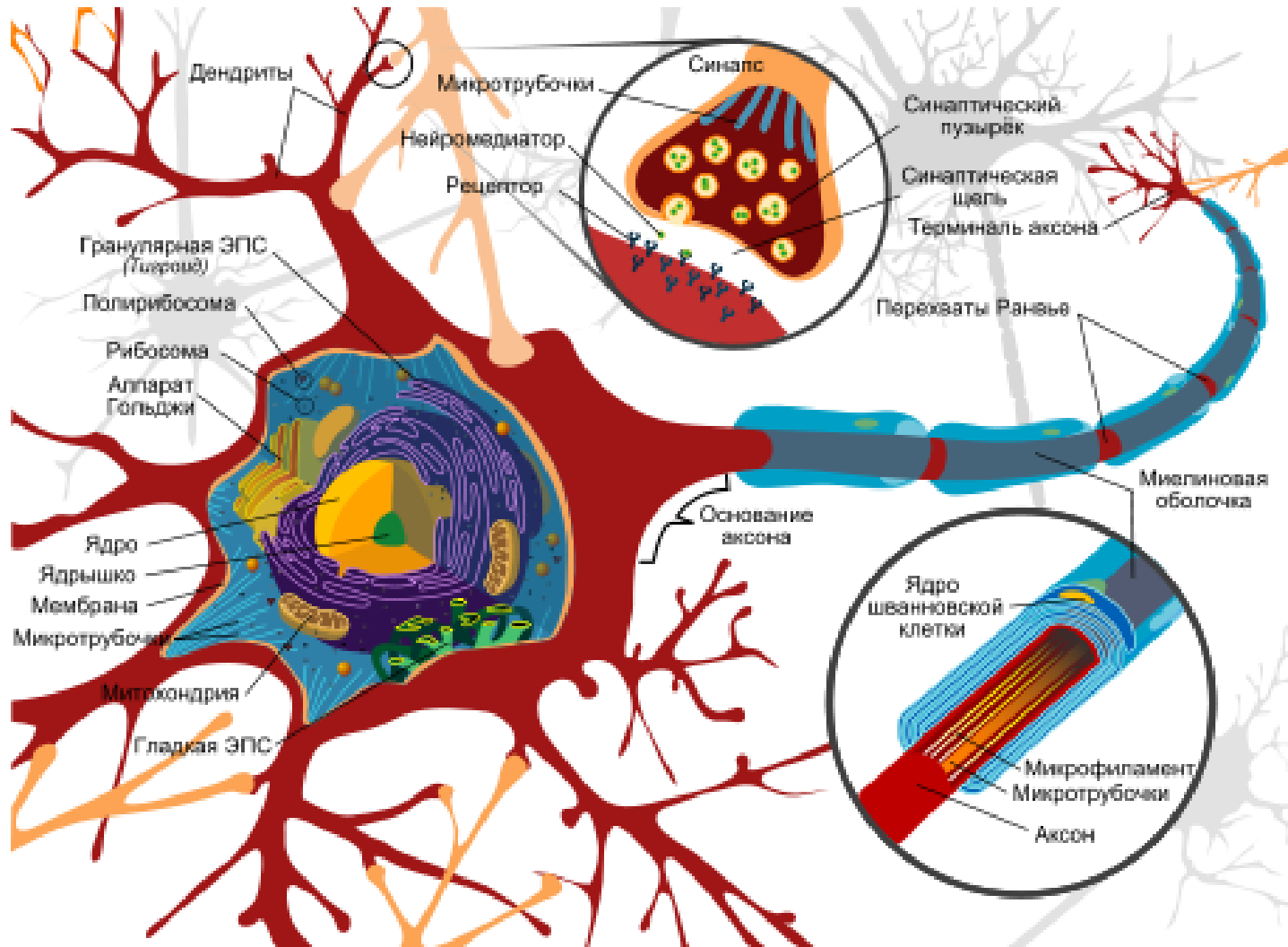


Эволюция мозга



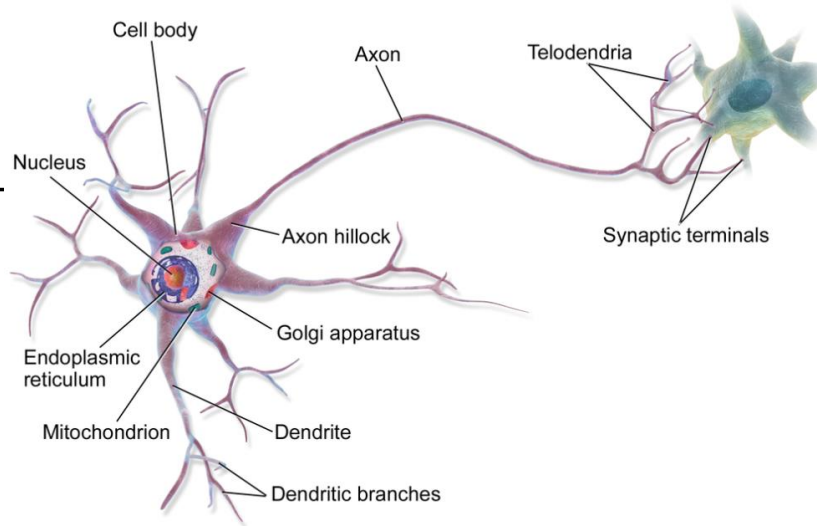
Нейрон

- электрически возбудимая клетка, которая обрабатывает, хранит и передает информацию с помощью электрических и химических сигналов.



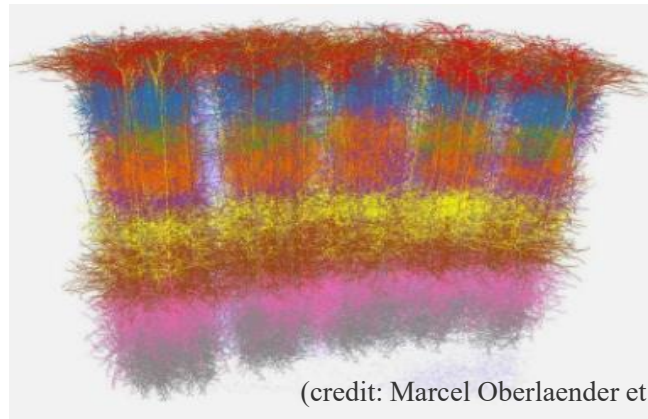
Каждую секунду мозг решает много задач : *восприятие*, контроль органов, движение, обучение, *разговор*, ...

1. В мозге ~80 миллиардов нейронов. Каждый нейрон - небольшой компьютер, распознающий множество паттернов.



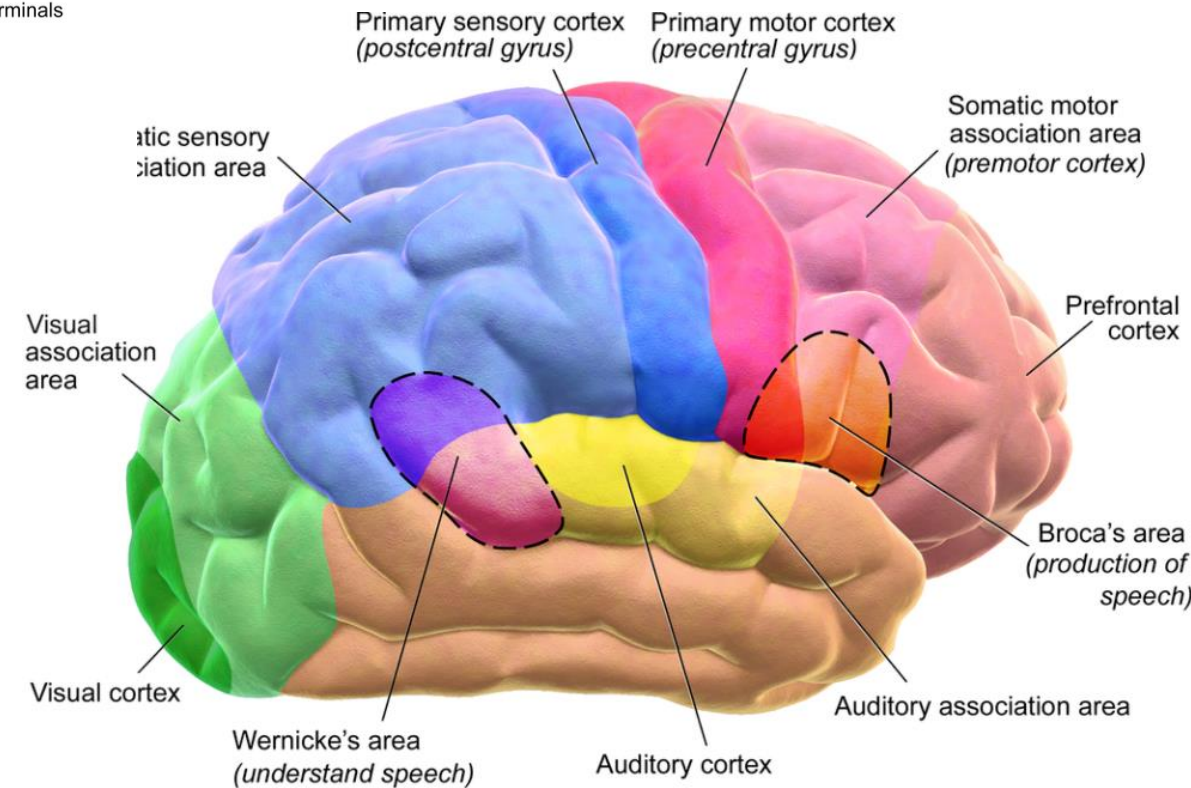
(credit: [BruceBlais](#))

2. Нейроны объединяются в кортикальные колонки – каждая решает свою задачу распознавания и передачи информации.



(credit: Marcel Oberlaender et al.)

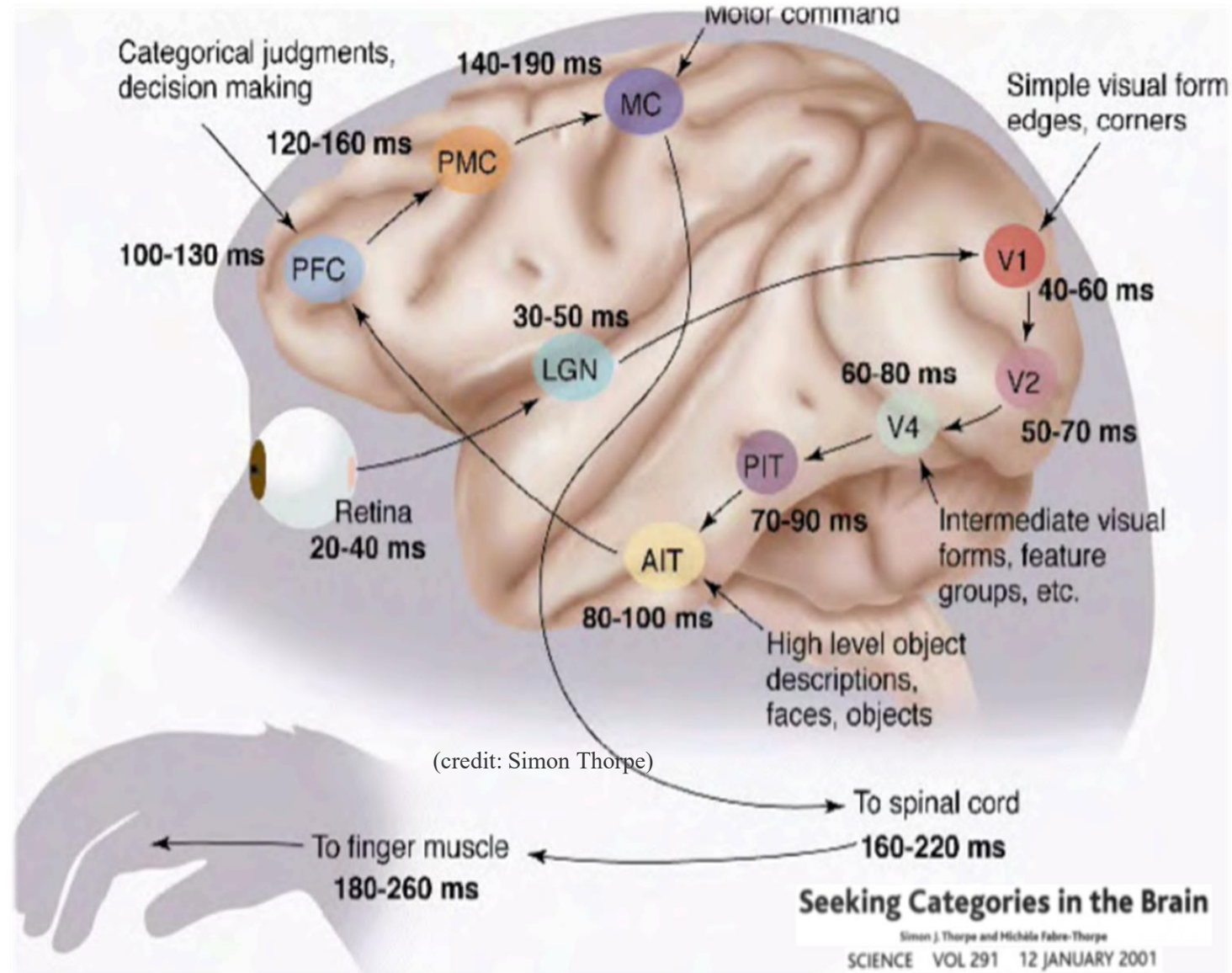
3. Мозг разделен на функциональные поля (visual, hearing, emotions, motion...).



Credit: Blausen.com staff (2014). "[Medical gallery of Blausen Medical 2014](#)".

Зрительная система

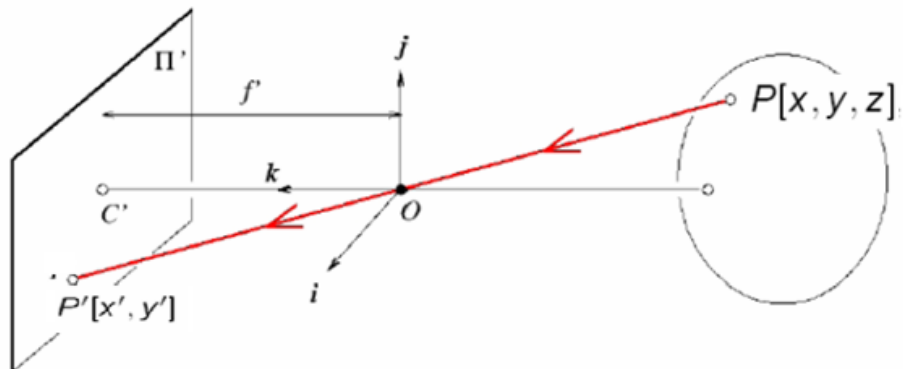
- Иерархическая организация полей нейронов.
- Каждое поле можно рассматривать как слой нейронов, который извлекает все более сложную информацию из потока данных.
- V1 – первичная визуальная кора: по сигналам с сетчатки распознает цветовые пятна, особые области и изменения интенсивности в разных направлениях.
- V2, V3, V4 : бинакулярное зрение, простые геометрические формы, движение.
- Inferior Temporal Cortex (PIT, AIT areas) - is a final stage in the ventral cortical visual system. Нейроны “узнают” лица, эмоции, руки, жесты и другие сложные объекты.



Эволюция камер



Модель камеры-обскуры используется до сих пор в компьютерном зрении



$$x' = f' \frac{x}{z}$$

$$y' = f' \frac{y}{z}$$

Не только RGB

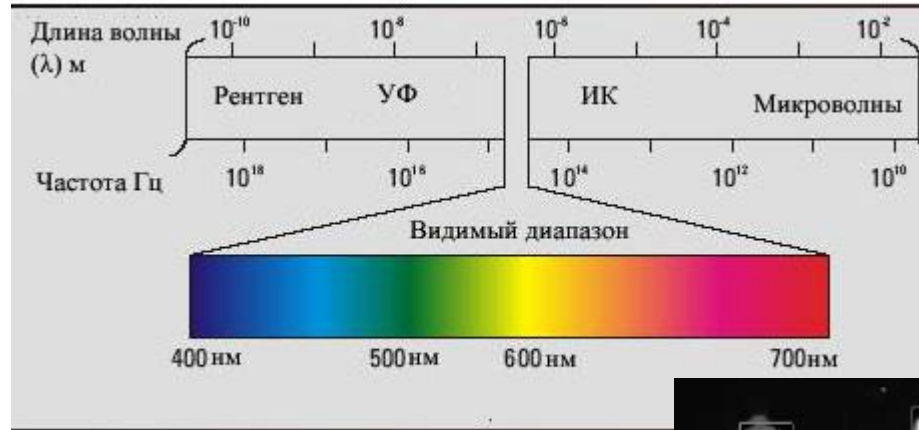


рис. 1

- Инфракрасное зрение
- Термальное зрение
- Рентгеновское зрение
- Терагерцовое зрение
- 3D зрение, стерео



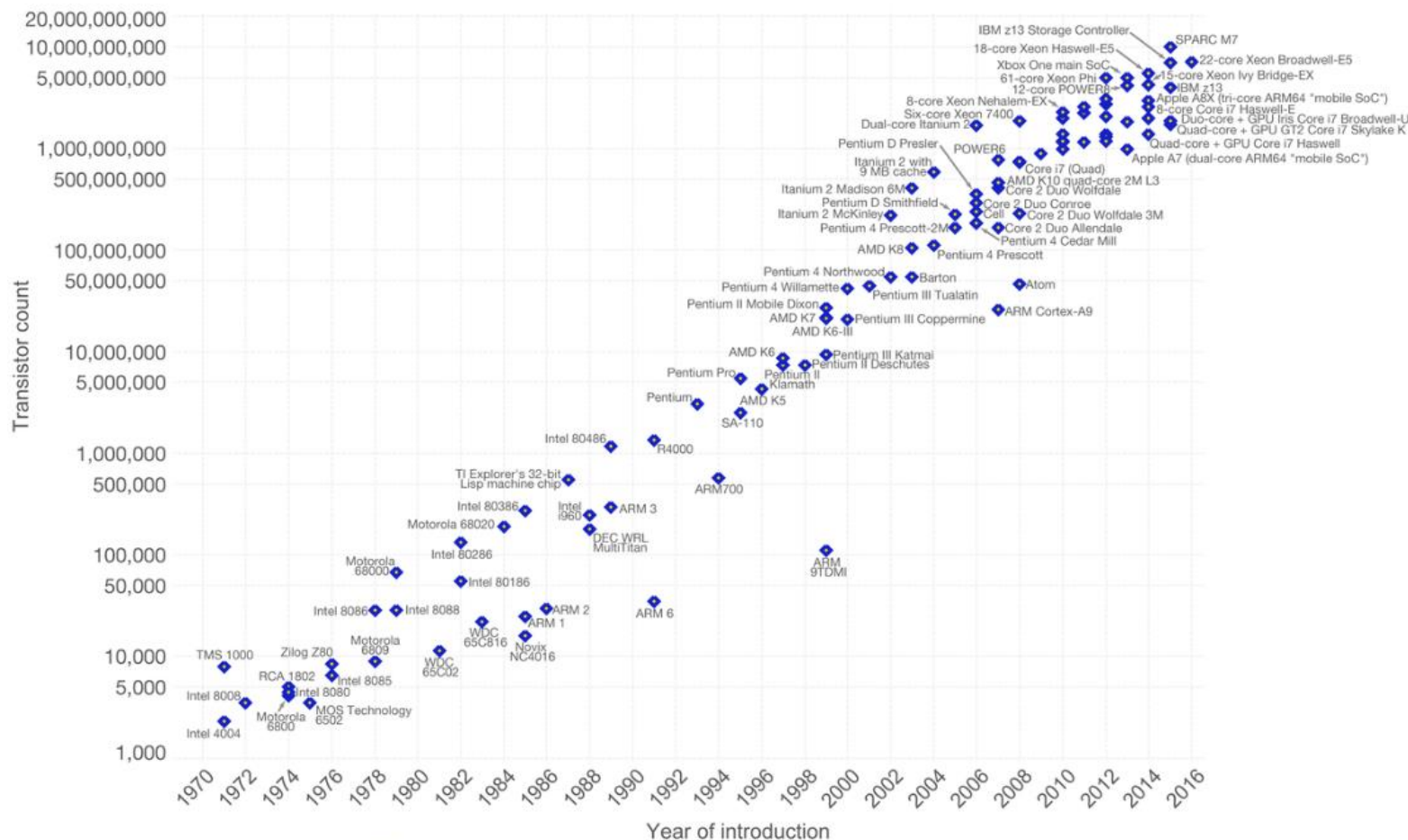
Эволюция вычислительных мощностей

Закон Мура - количество транзисторов, размещаемых на кристалле интегральной схемы, удваивается каждые 24 месяца.

Moore's Law – The number of transistors on integrated circuit chips (1971-2016)


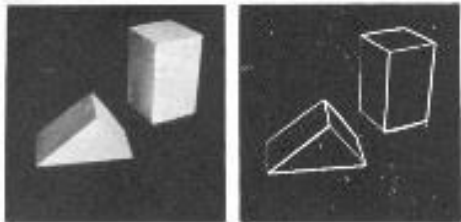


Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are strongly linked to Moore's law.



Как все начиналось: 196х

Machine Perception Of Three-Dimensional Solids
Massachusetts Institued of Technology
Lincoln Laboratory
22 May 1963
Lawrence G. Roberts



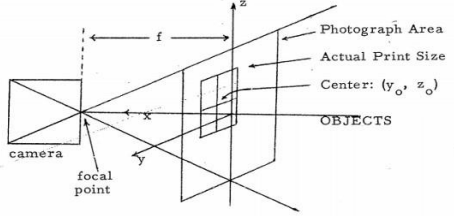


Figure 1: Camera Transformation

Photograph Area
Actual Print Size
Center: (y₀, z₀)
OBJECTS
camera
focal point

$ax + by = c$

$a = \sum x$	$\sum y^2$	$-\sum y$	$\sum xy$
$b = \sum y$	$\sum x^2$	$-\sum x$	$\sum xy$
$c = \sum x^2$	$\sum y^2$	$-\sum xy$	$\sum xy$

1966:
1st chatbot

```
=====
EEEEEEE L      IIIIII ZZZZZZZ AAA
E      L      I      Z      A  A
EEEEEE  L      I      Z      A  A
E      L      I      Z      AAAAAAA
EEEEEEE LLLLLLL IIIIII ZZZZZZ A  A
=====
ELIZA > (PLEASE TYPE IN ALL CAPS) WHAT'S YOUR NAME DEAR ?
AMIT
ELIZA > HI, I'M ELIZA. WHAT DO YOU WANT TO TALK ABOUT ?
AMIT > CAN I TALK ABOUT MY PROBLEM ?
ELIZA > SURE...!
AMIT > DUE TO ON GOING EXAMS IN COLLEGE, I HAVE A LOT OF STRESS.
ELIZA > PLEASE RELAX A LITTLE. SLEEP WELL.
AMIT > THANKS FOR YOUR ADVICE.
ELIZA > NO MENTION.
AMIT > BYE
ELIZA > BYE AND KEEP IN TOUCH...
=====
```

1959: An apparatus that allowed transforming images into grids of numbers (176x176)

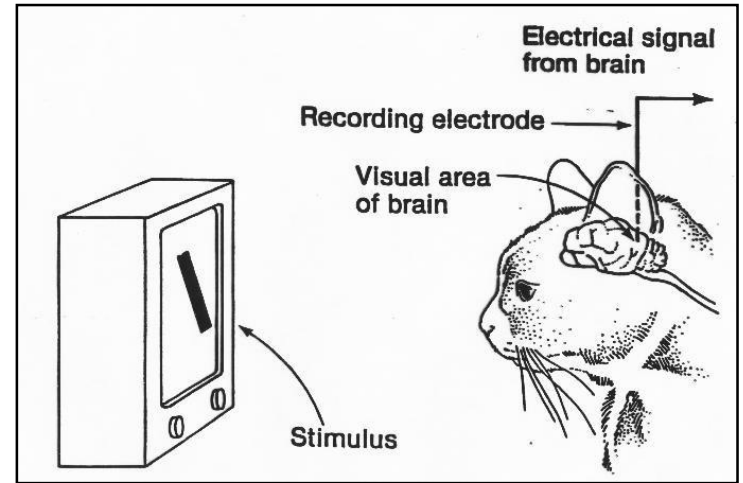


1966:IBM's speech recognition

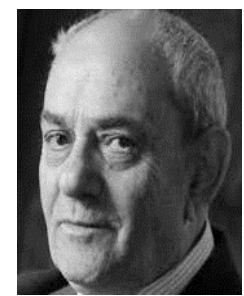
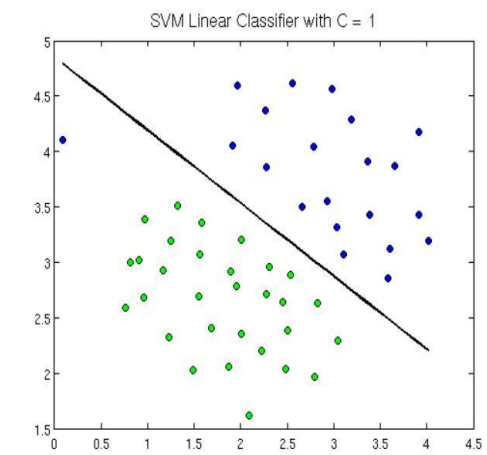


1965: Herbert Simon predicts that **"Machines will be capable, within 20 years, of doing any work a man can do"**
AI became an academic discipline.

1959: Receptive fields of single neurons in the cat's striate cortex



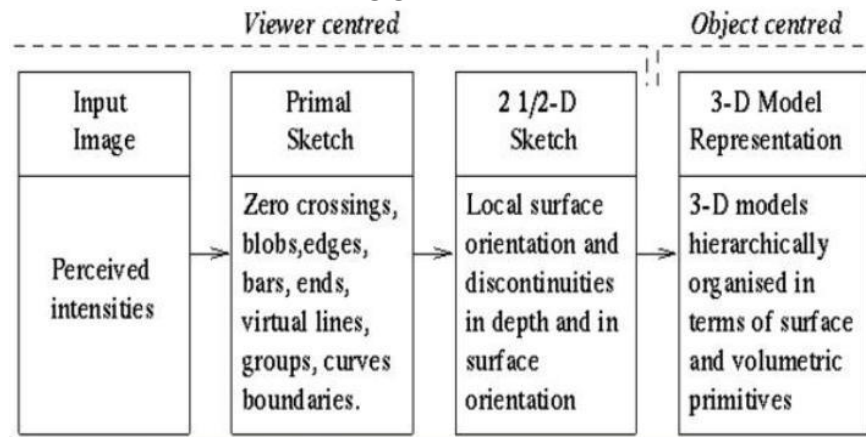
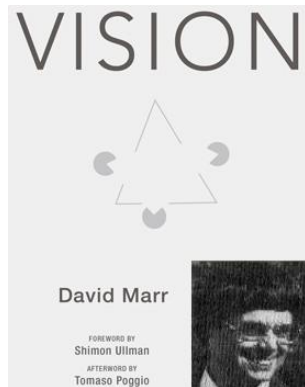
1963: Vladimir Vapnik: Support Vector Networks



197x:

1971: The "Stanford cart" – автономная машина, объезжающая препятствия

1979: David Marr's and Tomaso Poggio's "2 1/2 sketch"



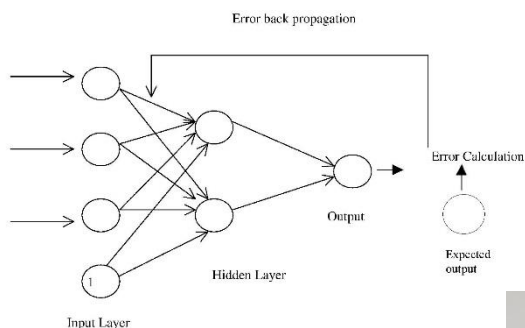
http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/GOMES1/marr.html

1974: Paul Werbos' backpropagation algorithm for neural networks

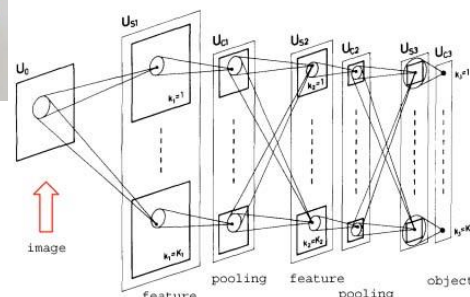


BEYOND REGRESSION:
NEW TOOLS FOR PREDICTION AND ANALYSIS
IN THE BEHAVIORAL SCIENCES

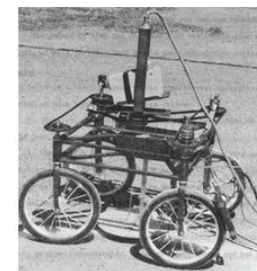
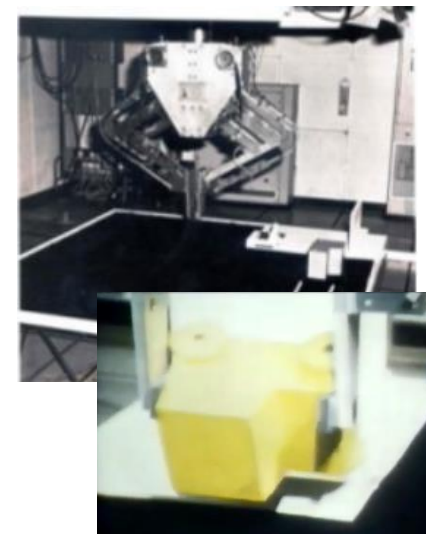
Harvard University
Cambridge, Massachusetts
August, 1974



1979: Kunihiro Fukushima's convolutional neural network ("Neocognitron - A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position")



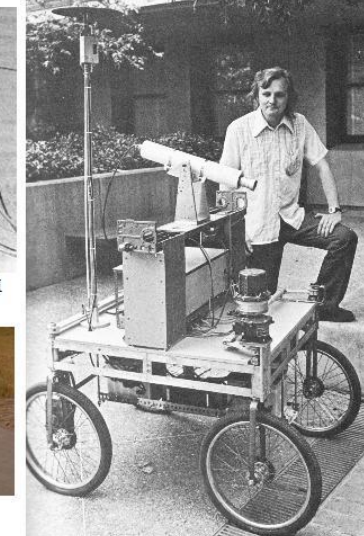
1973: Freddy – робот с машинным зрением



Jim Adams' "cart" in 1961



Rodney Schmidt's "cart" in 1971



Hans Moravec's "cart" in 1977

1974, [Kurzweil Computer Products](#) offered their first optical character recognition (OCR)



The growing criticism of the AI and computer vision technology resulted in "AI winter".

198x:

1981: famous [Lukas-Kanade optical flow algorithm](#)



1986: classical [Canny edge detector](#)



1987: EigenFace algorithm



1989: Yann LeCun's convolutional neural network for handwritten-digit recognition (LeNet-1)

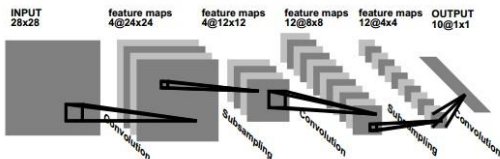


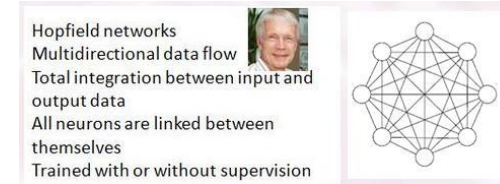
Figure 1: Architecture of LeNet 1



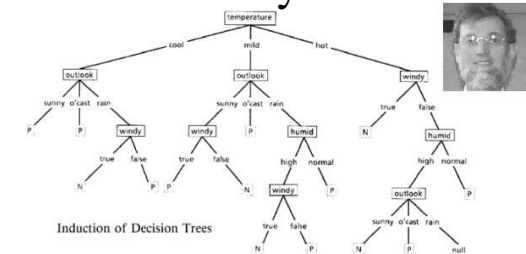
1981: Automatrix introduces the first commercial robot with a vision system



1982: John Hopfield describes a new generation of neural networks, based on recurrence



1985: Ross Quinlan's ID3 for decision trees analysis



1983: Scott Kirkpatrick's simulated annealing



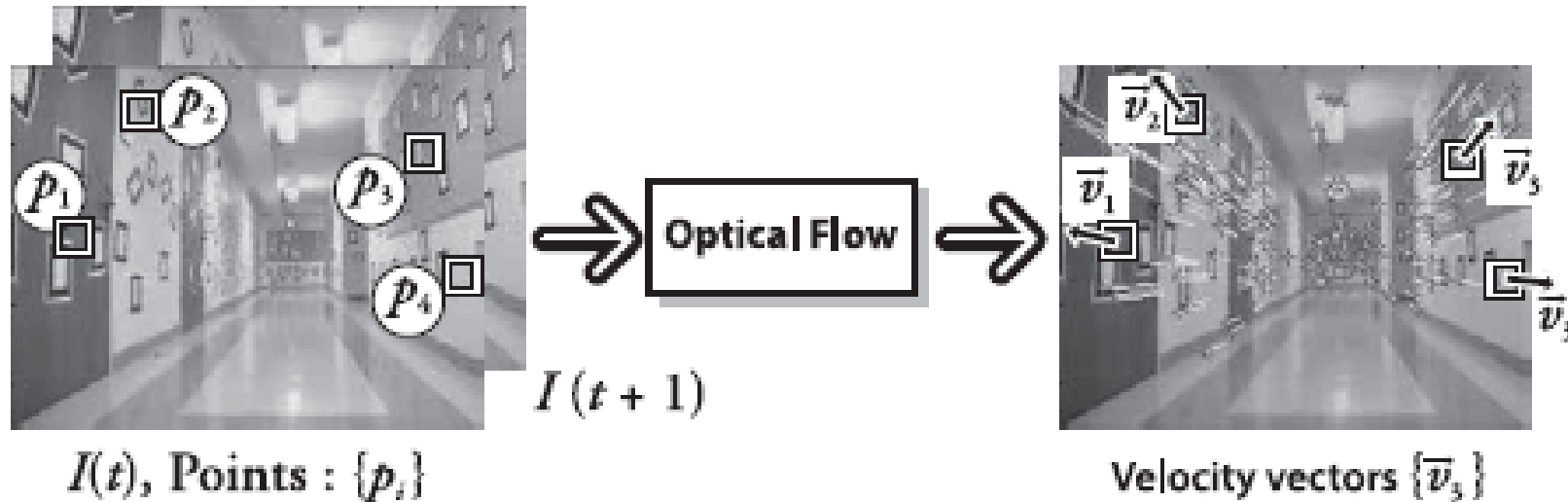
1988: Dean Pomerleau's self-driving vehicle ALVINN



Оптический поток

$$\sum_{x=p_x-\omega_x}^{p_x+\omega_x} \sum_{y=p_y-\omega_y}^{p_y+\omega_y} (A(x,y) - B(x+\nu_x, y+\nu_y))^2 \longrightarrow \min$$

$$\bar{\nu} = [\bar{\nu}_x \quad \bar{\nu}_y]^T$$

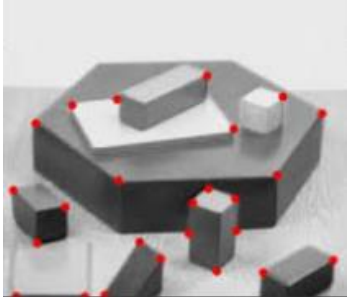


picture from:
Gary Bradski, Adrian Kaehler, "Learning OpenCV:
Computer Vision with the OpenCV Library"

$$\bar{\nu}_{\text{opt}} = G^{-1} \bar{b}, \quad G \doteq \sum_{x=p_x-\omega_x}^{p_x+\omega_x} \sum_{y=p_y-\omega_y}^{p_y+\omega_y} \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad \bar{b} \doteq \sum_{x=p_x-\omega_x}^{p_x+\omega_x} \sum_{y=p_y-\omega_y}^{p_y+\omega_y} \begin{bmatrix} \delta I I_x \\ \delta I I_y \end{bmatrix}$$

199x:

1994: Shi and C. Tomasi.
Good Features to Track.

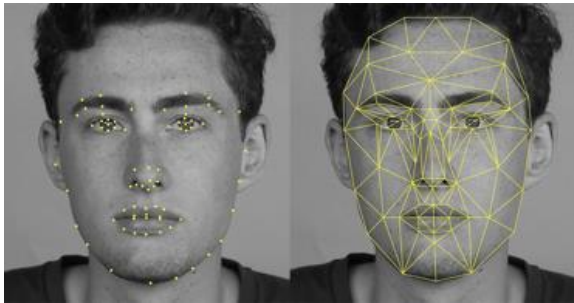


1999: David Lowe, Object recognition
from local scale-invariant features

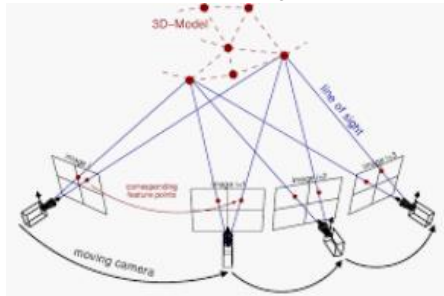


OpenCV (Open Source Computer Vision) -
a popular computer vision library started
by Intel in 1999.

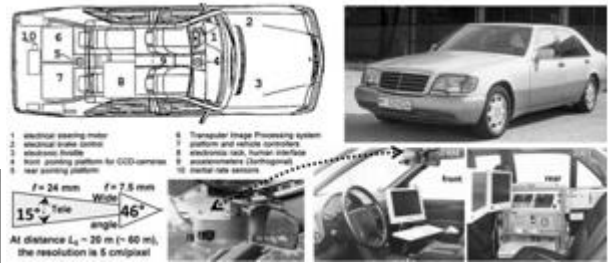
1995: Cootes et al.
Active Shape Modelling



1999: Triggs et al.,
"Bundle Adjustment —
A Modern Synthesis"



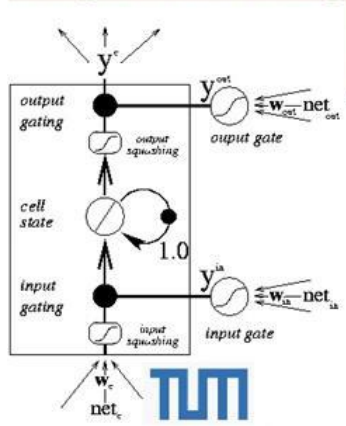
1994: Ernst Dickmanns' self-driving car drives more than 1,000
kms near the airport Charles-de-Gaulle in Paris



1997: IBM's "Deep Blue" chess
machine beats the world's chess
champion, Garry Kasparov



Long Short-Term Memory



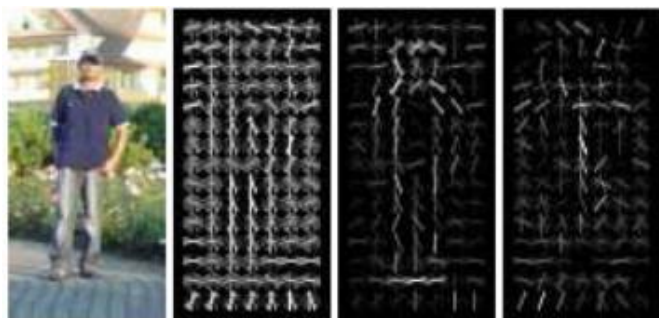
1997: Jeürgen Schmidhuber's
and Sepp Hochreiter's Long
Short Term Memory (LSTM)
model

200x:

2001: Viola/Jones face detector



2005, Dalal, Triggs,
[Histograms of Oriented
Gradients for Human
Detection](#)



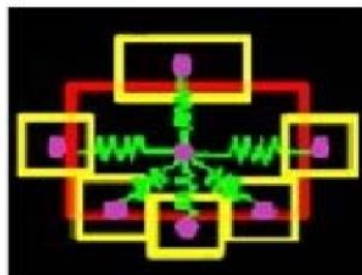
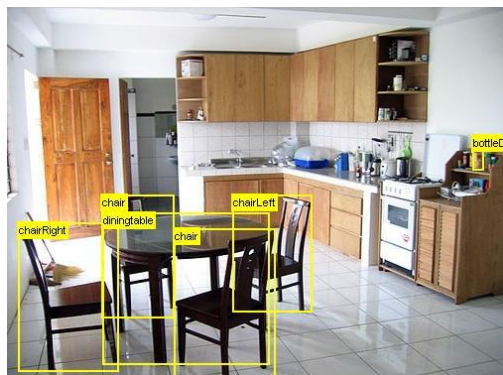
2005: driverless car Stanley wins
DARPA's Grand Challenge



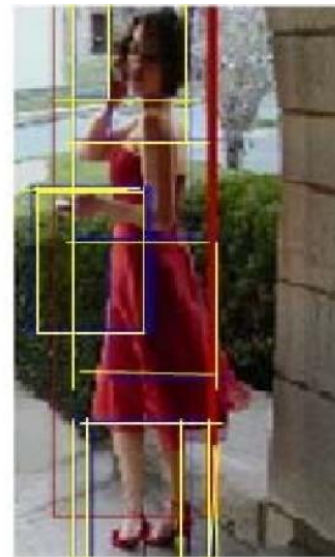
2006: Scott Hassan founds robot startup Willow Garage



2007, The PASCAL Visual Object Classes Challenge 2007



2009: Felzenszwalb et al,
[Deformable Part Model](#)

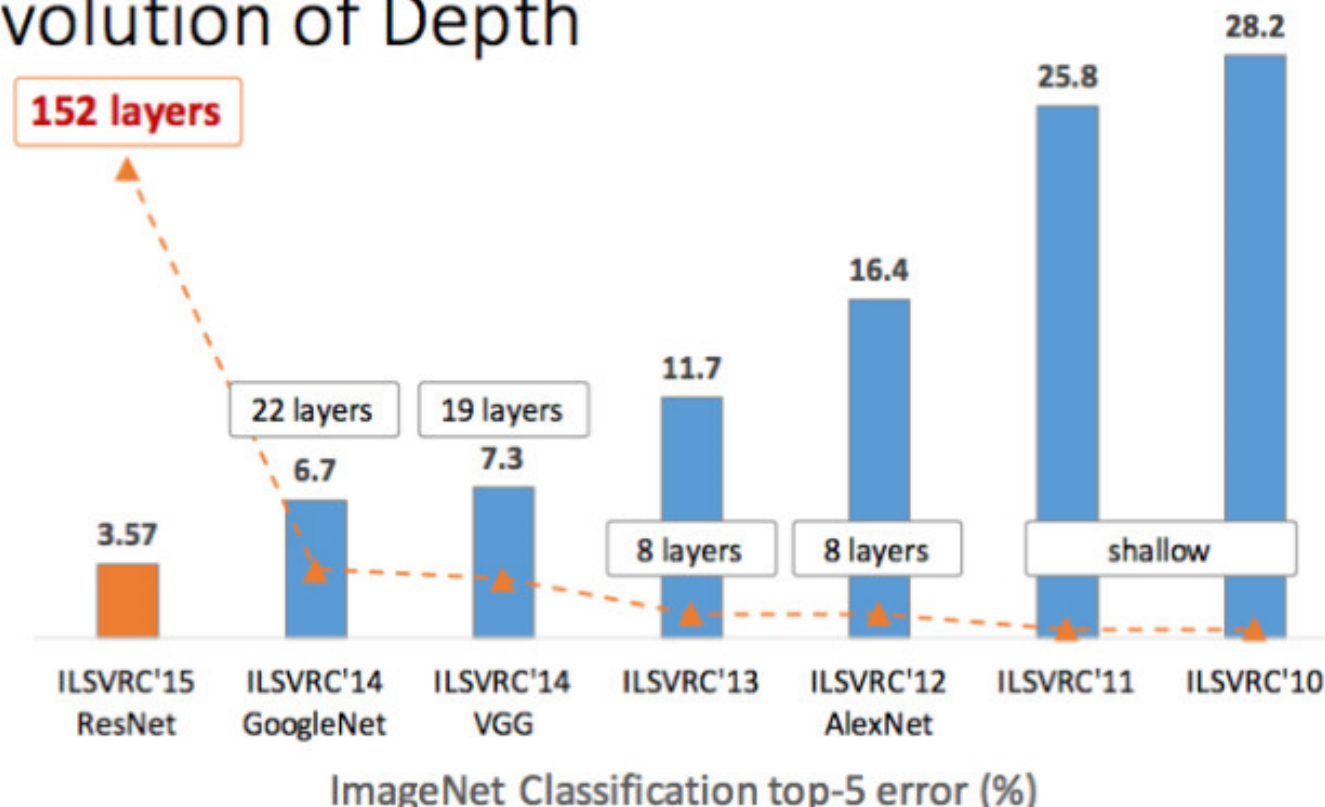


201x

- 2010: The ImageNet Large Scale Visual Recognition Competition
- 2012 when [AlexNet won ImageNet](#)
- 2014, Goodfellow et al, **Generative Adversarial Networks**



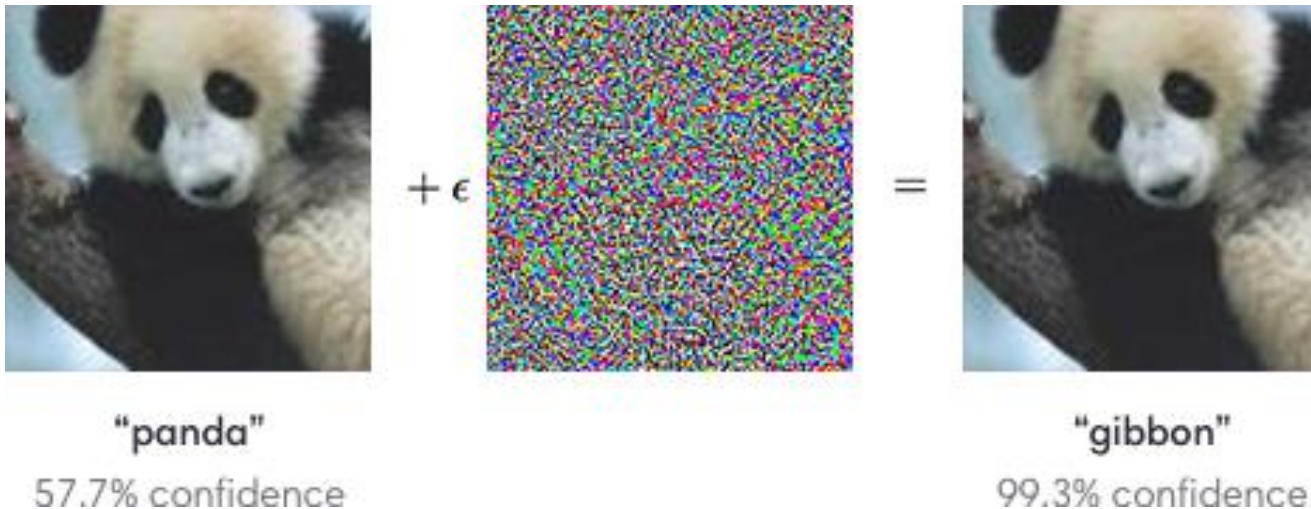
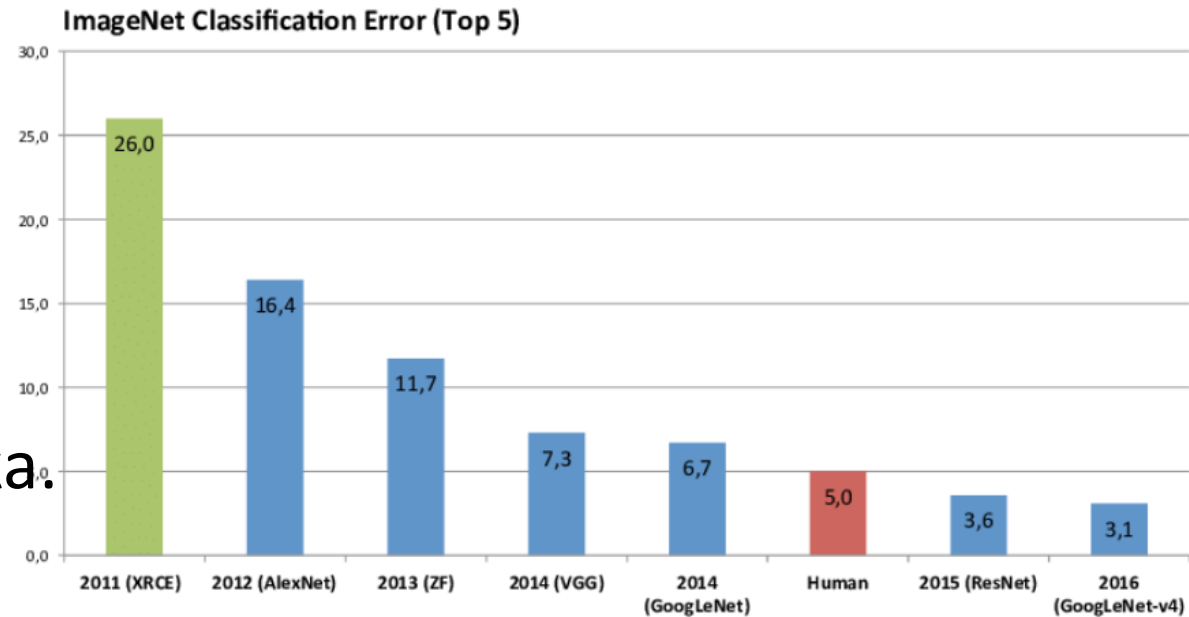
Revolution of Depth



Сегодня

Алгоритмы компьютерного зрения:

- Лучше человека классифицируют.
- Лучше человека распознают человека.
- Рисуют и подделывают (DeepFake).
- Neural Architecture Search: сети делают сети.
- НО!



Завтра

- ?

Итак...

- Появление “полноценного” зрения в природе возможно спровоцировало эволюционный взрыв. То же будет и в технике?
- Технологическая эволюция зрения ускоряется с каждым годом. Эволюция алгоритмов: от оптимизма (196х) к еще большему оптимизму (201х).
- Компьютерное зрение уже решает некоторые задачи лучше человека. В тоже время таких задач пока не так много. Да и решения эти не устойчивы к шуму и к исключительным случаям.