Neural Network Theory and Applications
神经网络理论及应用

Bao-Liang Lu（吕宝粮）
Dept. of Computer Science & Engineering
Shanghai Jiao Tong University
Office: 3-431 Dianxin Bld, Minhang
Tel: 3420-5422
bllu@sjtu.edu.cn

AlphaGo掀起了新一轮AI浪潮！

- 2016年1月，Google在Nature上发文，表示他们的AlphaGo系统在正式比赛中打败了欧洲围棋冠军。
- 2016年3月，AlphaGo成功挑战韩国围棋九段李世石，比分为4:1
- 2017年年初Master横扫人类棋手！
- AlphaGO:强化学习+深度学习+搜索技术

AlphaGo Zero

- **AlphaGo Zero is a version of AlphaGo, which is created without using data from human games and stronger than any previous version.**
- **AlphaGo's Zero is trained without datasets derived from human experts. All of the training data are generated by playing games against itself.**

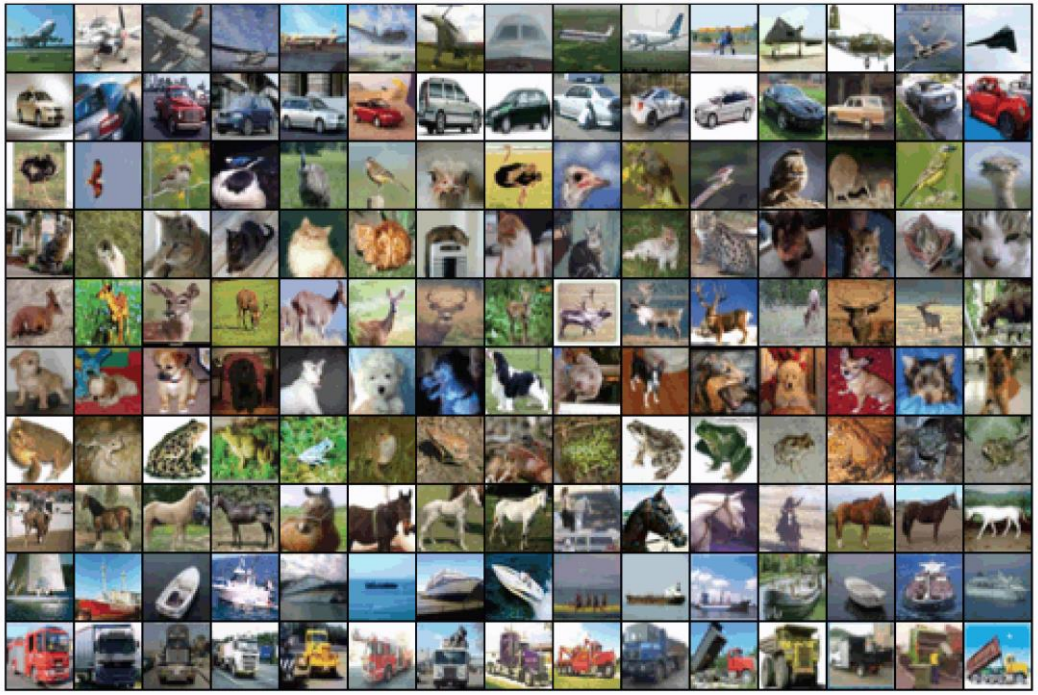
2017年10月19日凌晨，在国际学术期刊《自然》（Nature）上发表的一篇研究论文中，谷歌下属公司Deepmind报告新版程序AlphaGo Zero：从空白状态学起，在无任何人类输入的条件下，它能够迅速自学围棋，并以100:0的战绩击败“前辈”。

ImageNet

- **Information**
  - Total number of non-empty synsets: 21841
  - Total number of images: 14,197,122
  - Number of images with bounding box annotations: 1,034,908
  - Number of synsets with SIFT features: 1000
  - Number of images with SIFT features: 1.2 million

- **The contest contains 1000 different categories, including animals, fruits, vegetables and so on.**
  - Training set: 1.2 millon
  - Test set: 100 k
  - Validation: 50 k

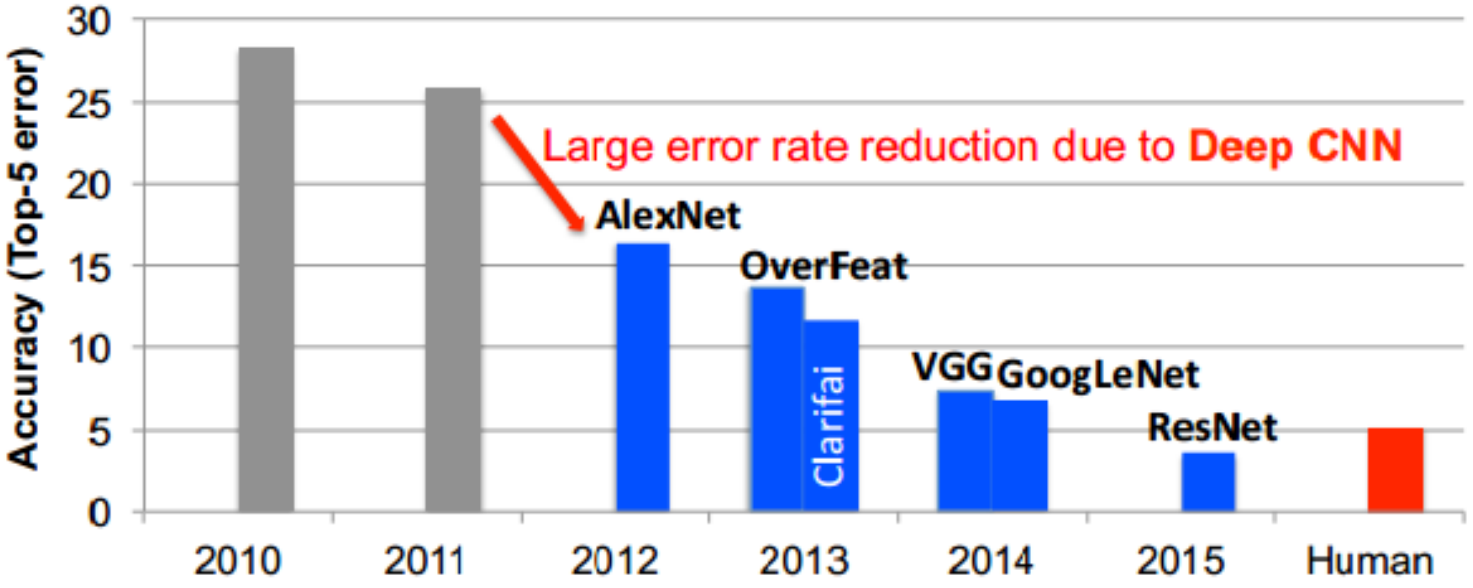
ImageNet



ImageNet

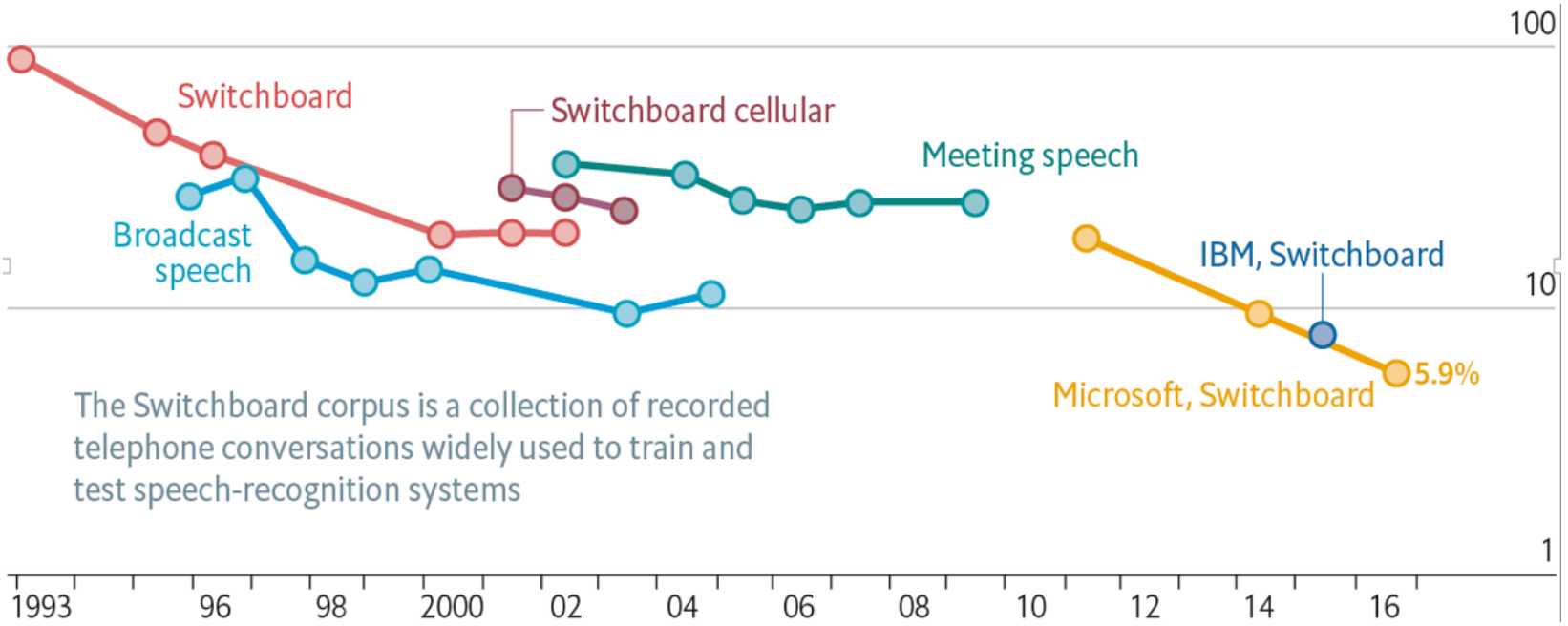
ImageNet and Deep Learning

- **The top-5 error rate has reduced drastically when the deep learning methods were adopted to solve ImageNet classification problems. The error rates are smaller than human's judge in recent methods.**



Speech Recognition and Deep Learning

- **The error rate of speech recognition has a high value before 2010. The application of deep learning in this field makes the error rate drastically reduced.**

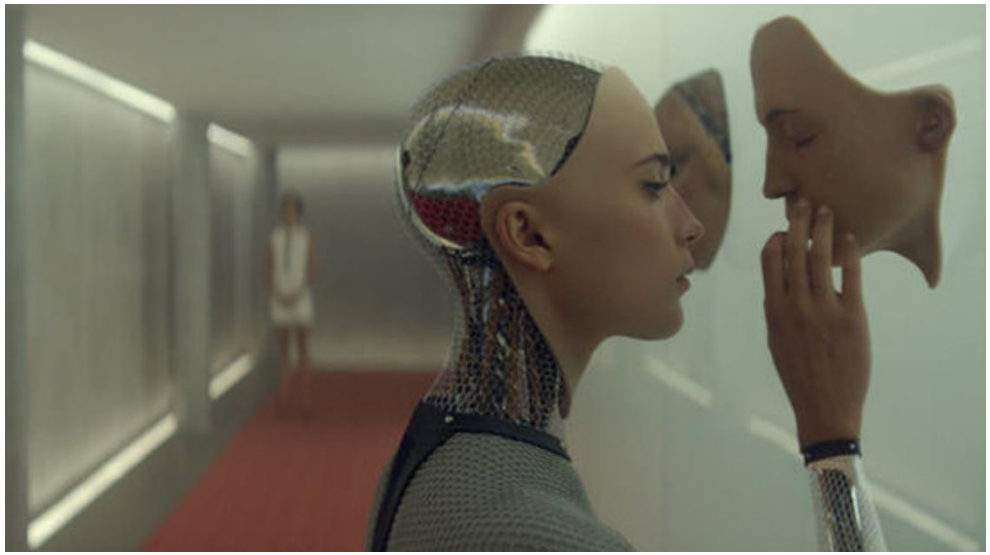


Artificial General Intelligence

通用人工智能
强人工智能

Ethics of artificial intelligence

- With the rapid developments of the AI-based systems, large amount of questions related to the social impact, governance, and ethical implications of these technologies have emerged .
- The Berkman Klein Center (Harvard) and the MIT Media Lab have set up the Ethics and Governance of Artificial Intelligence Fund to solve the ethics problems caused by AI.



Goals for the Course

- Learn the theories, algorithms, methods and foundational ideas of NN
- Prepare to apply NN
- Prepare to do research in NN and related fields
- Learn some new ways of thinking about AI, machine learning, and Intelligent information processing
  - The biological perspective
  - The systematical perspective
  - The skeptical perspective

Course Mechanics

- Programming exercises : each student will do 6 of these, including one of own devising (in consultation with instructor).
- Grading: 40% programming exercises; 40% final project; 20% quiz and class participation.
- Deadline of final Project: June 30, 2018

Instructor and Slides

- My Ph.D. student：Wei Liu（刘伟），Yun Luo（罗赟）
- Email：liuwei@liujr.com；angeleader2012@163.com
- Send email to Wei Liu:
  - Your name
  - Your student ID
  - Your email
- You can download the slides from the following FTP:

  ftp://bcmi.sjtu.edu.cn:2122/

  User name: nnet

  Password: nnet

Outline of Lecture

- Introduction
- Perceptron, Multilayer Perceptron
- Min-Max Modular Network
- Support Vector Machine
- Unsupervised Learning: Self-Organization Map
- Auto-encoder
- Transfer Learning
- Convolutional Neural Network (CNN)
- Long and Short-Term Memory (LSTM)
- Generative Adversarial Networks (GAN)
- Reinforcement Learning

世界主要国家脑计划简介

大脑：通用人工智能的唯一参照物

- 物理大脑：
  - ~1.3升，~1.5公斤，占体重2%
  - 功耗~20W，占全身20%
- 神经大脑：
  - 神经元数约1000亿（860亿）
  - 1~2M个功能柱（6层），每柱约1万神经元，人脑突触总数约100T
  - 每个神经元通过数千至上万突触与其他神经元相联
  - 神经元典型发放频率不超过100Hz
  - “除人脑外，没有任何一个自然或者人工系统能够具有对新环境与新挑战的自适应能力、新信息与新技能的自动获取能力、在复杂环境下进行有效决策并稳定工作直至几十年的能力。没有任何系统能够在多处损伤的情况下保持像人脑一样好的鲁棒性，在处理复杂任务的同时，没有任何系统能媲美人脑的低功耗性”
    ——HBP建议书

日本脑科学研究计划（1996-2016）

- 理解脑 （Understanding the Brain)领域
  - 神经细胞机能部（4）
  - 神经网络机理部（4）（伊藤正男教授）
  - 认知脑科学部（2）
- 保护脑 (Protecting the Brain)领域
  - 发生与分子化学研究部（2）
  - 病因遗传基因研究部（3）
- 创造脑 (Creating the Brain)领域
  - 脑型器件与脑道研究部（2）（松本元教授）
  - 脑型信息系统研究部（3）（甘利俊一教授）
- 先端技术开发中心（2）

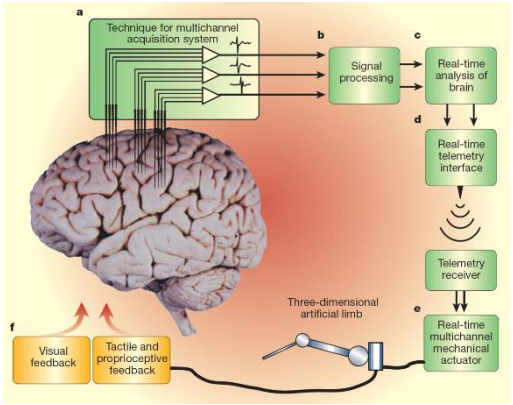
（1997年当时实验室22个；鼎盛时期发展到近50个实验室）

2013年欧美的脑科学研究计划

- 欧盟未来新兴旗舰技术项目（21>6>2)
  - 石墨烯；人脑工程
- 人脑工程
  - 2013年9月启动；10年将投入10亿欧元
  - 瑞士洛桑联邦理工学院教授马克拉姆牵头
  - 87个世界各地研究团队承担
  - 2023年前完成
  - 目标：在治疗帕金森氏症、研发超级计算机和智能机器人等领域取得重大突破
- 奥巴马政府实施了一项为期10年的科学研究，检查人脑的活动，并绘制一张全面的人脑活动图(Brain Activity Map)，希望能在大脑研究领域做出类似人类基因组计划 (Human Genome Project)为遗传学所做的贡献（纽约时报）

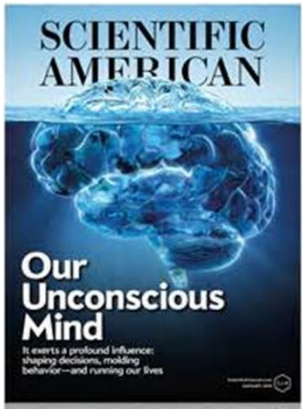
Brain Science and Technology Institute (Nov. 19, 2015)

- Three Centers
  - Mental heath
  - Brain-inspired Computing
  - Brain Diseases
- Brain-inspired Computing
  - Brain-inspired computing theory
  - Brain-computer interaction
  - Brain-inspired chip and device
- Brain computer Interaction
  - Brain function rehabilitation
  - Diagnosis of mental diseases such as attention deficit hyperactivity disorder (ADHD)
  - Drowsy driving prediction
  - Affective brain-computer interaction
- China Brain research Project (2016-2030) will started this year ?



科学美国人：关于人类未来的20个大问题

1. 人类的未来会超越地球吗？
2. 何时何地能发现外星生命？
3. 我们有可能理解意识的本质吗？
4. 每个人都能得到必要的医疗护理吗？
5. 脑科学会改变刑法吗？
6. 500年后人类还存在吗？
7. 我们离阻止核毁灭更近了吗？
8. 性会变得过时吗？
9. 人体所有器官都可以人造吗？
10. 我们能避免第六次大灭绝吗？



2016年9月1日

11. 我们能在不破坏地球的情况下喂饱所有人吗？
12. 我们能殖民外太空吗？
13. 我们能找到地球的“孪生兄弟”吗？
14. 我们能治愈阿尔茨海默氏症吗？
15. 我们能用可穿戴技术探测自己的情绪吗？
16. 我们能揭示暗物质的真相吗？
17. 我们能控制精神分裂症或自闭症等脑部疾病吗？
18. 技术能让我们在药物研发过程中摒弃动物实验吗？
19. 科学领域能够实现男女平等吗？
20. 我们能预测地震等自然灾难并提前发出警示吗？

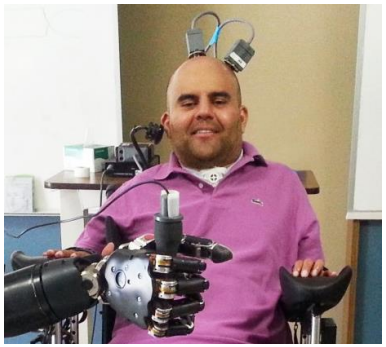
我们能用可穿戴技术探测自己的情绪吗？

- 回答者：MIT媒体实验室情感计算研究团队：Picard
- 情绪与生物化学和电信号有关，这些信号会到达我们体内的每个器官，例如，压力影响我们的身体和心理健康。可穿戴技术让我们能量化这些信号在较长时间内的模式。在未来十年，可穿戴设备可以为我们的健康提供个性化的预报。比如，基于用户最近的压力、睡眠、社交情感活动，让我们的健康和幸福指数在未来一周提高80%。智能可穿戴设备也能发现一些我们可以改变的模式，从而提高我们的健康和幸福指数。比如，通过让佩戴者每晚睡9小时以上并维持目前较低的压力水平，在接下来的四天内，其发病的可能性将会减少60%。未来20年，可穿戴式设备和从其获得的分析数据，能显著降低精神和神经疾病的发病几率。

What is brain-inspired intelligence?

- AlphaGO has powerful logic intelligence, but it doesn't have any **Emotional Intelligence**!

- Unlike industrial robots, serve robots must have ability to interact with humans and have emotional intelligence.



- Brain-inspired Intelligence=**Logic Intelligence+ Emotional Intelligence**

Outline of Lecture 1

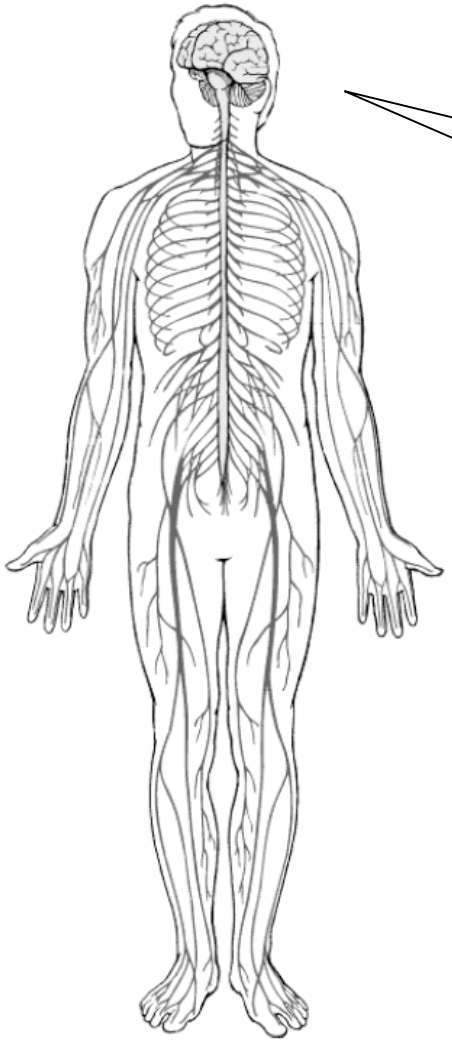
- Computer and Brain
- What is a neural network ?
- What can neural networks help ?
- What can neural networks do ?
- Biological neural networks
- Simple neuron model
- Neural network history
- Network architectures
- Learning paradigms
- Learning tasks
- Applications in my lab
- Recommended textbooks
- Neural network journals and conference

Comparing Brains with Digital Computers

	Computer	Human Brain
Computational units	1 CPU, 10^5 gates	10^{11} neurons
Storage units	10^9 bits RAM, 10^{10} bits disk	10^{11} neurons, 10^{14} synapses
Cycle time	10^{-8} sec	10^{-3}
Bandwidth	10^9 bit/sec	10^{14} bit/sec
Neuron updates/sec	10^5	10^{14}

A Very Short Introduction to Brain

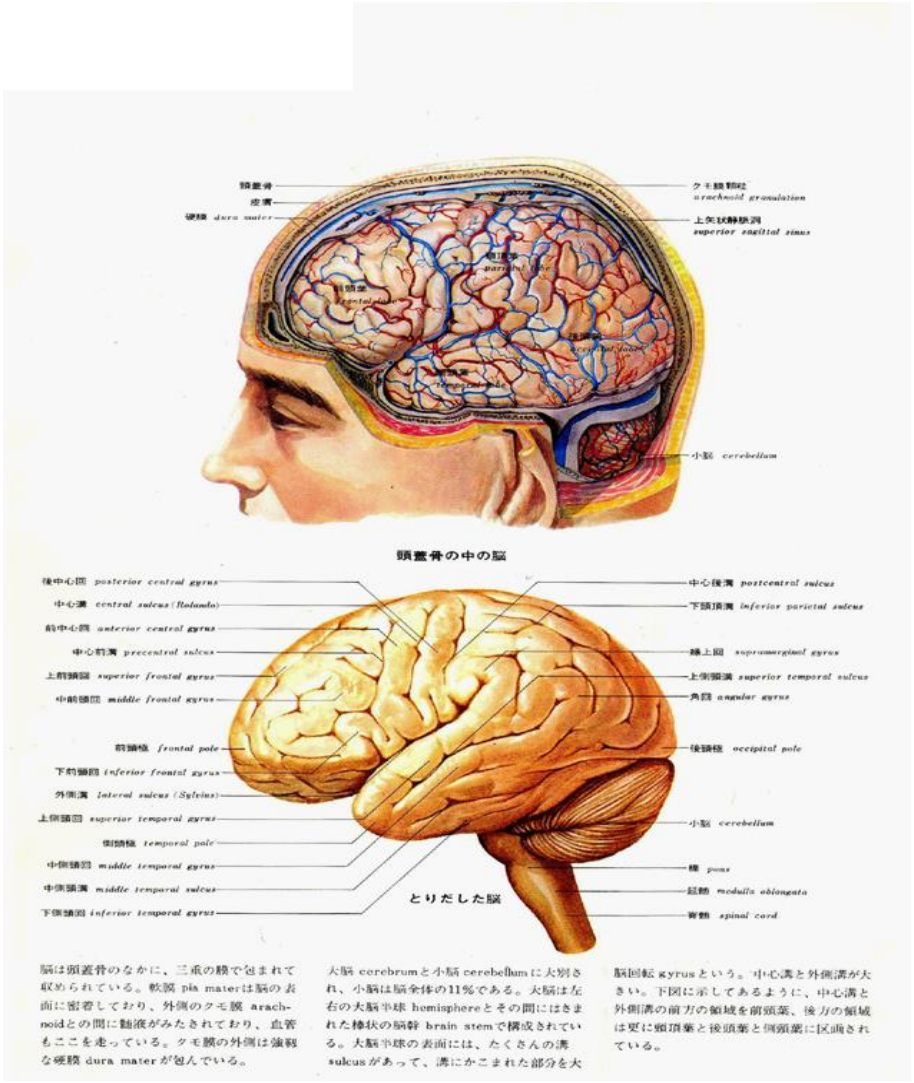
The Nervous System



Central nervous system
(Brain & spinal cord)

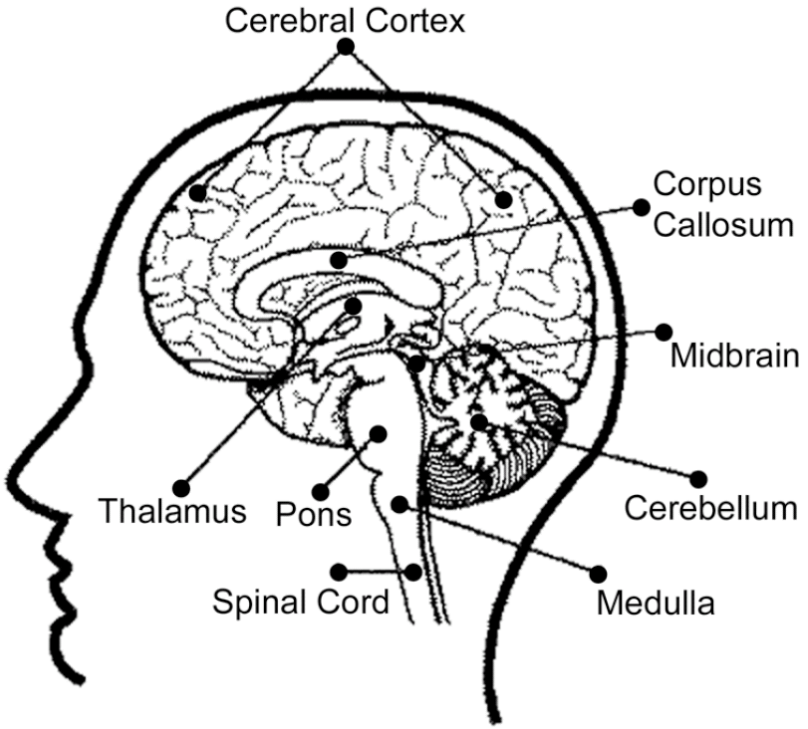
Peripheral nervous system

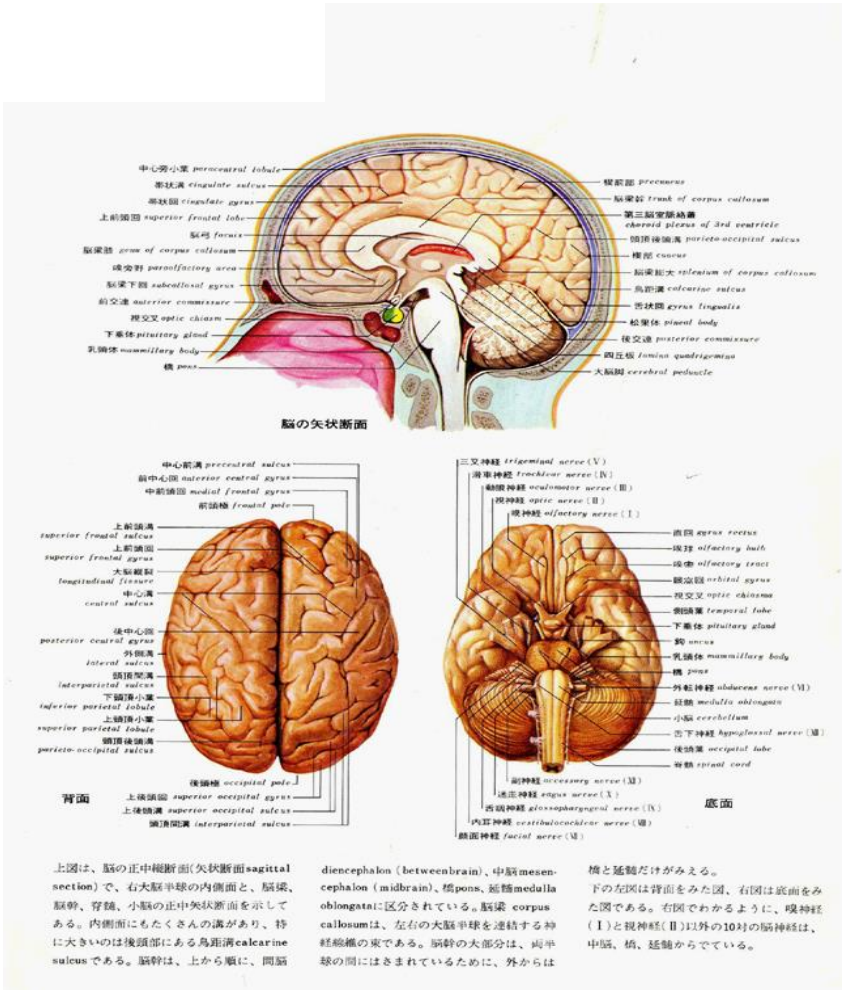
The Human Brain



- The brain contains about 100 billons neurons
- 10000 connections per neuron

Parts of the Human Brain



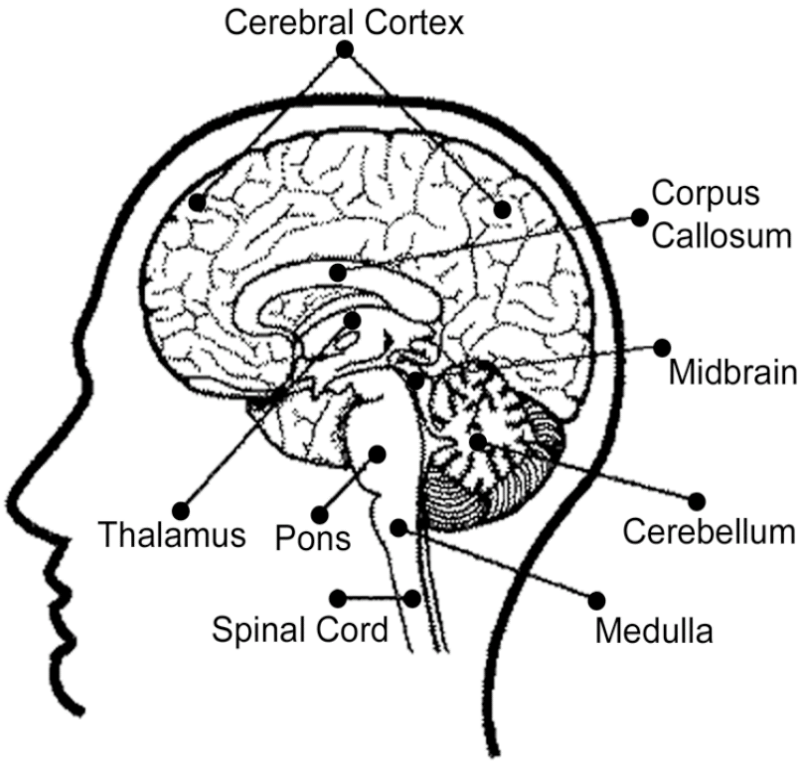


Cerebral Cortex 大脑皮层

- “Cortex” from the Latin for “bark” (of a tree)
- The cortex is a sheet of tissue that makes up the outer layer of the brain (2 to 6mm)
- The right and left sides of the cerebral cortex are connected by a thick band of nerve fibers
- FUNCTIONS: Thought, voluntary movement, language, reasoning, perception

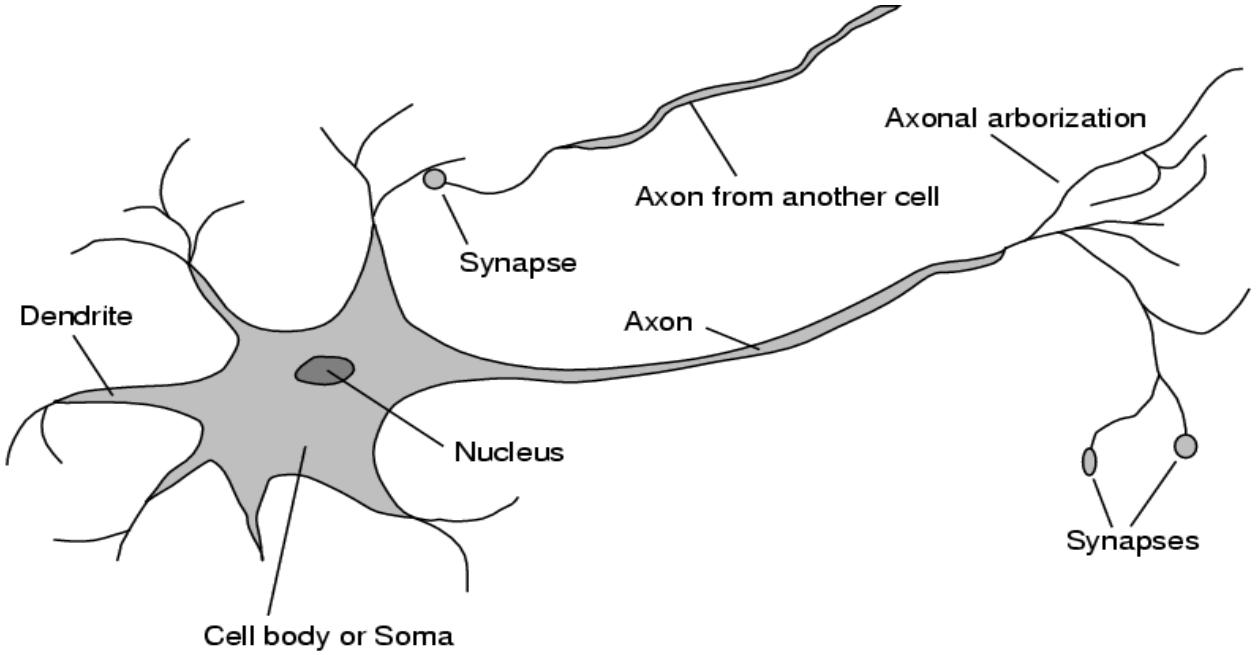
Cerebellum 小脑

- “Cerebellum” from the Latin for “little brain”
- FUNCTIONS: Movement, balance, posture



What is a Neuron ?

- Cells of the nervous system are called neurons (nerve cells)
- The human brain has about 100 billion neurons



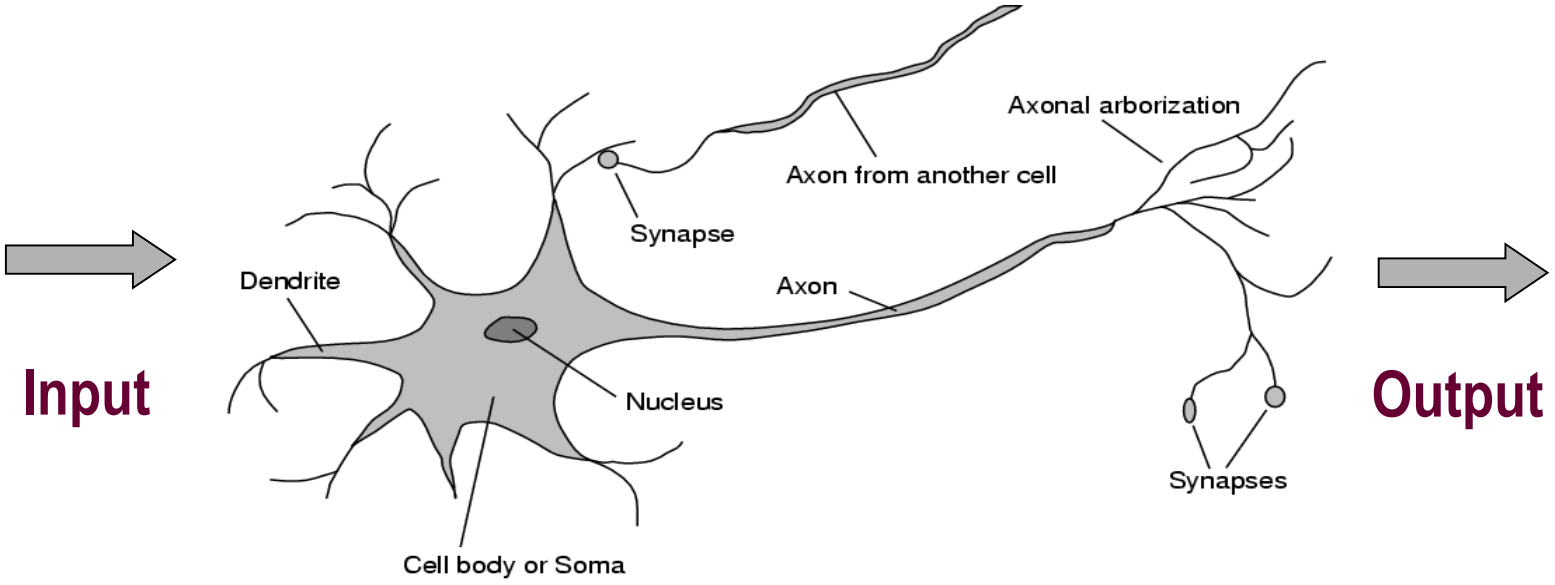
Axon : 轴突

Dendrite: 树突

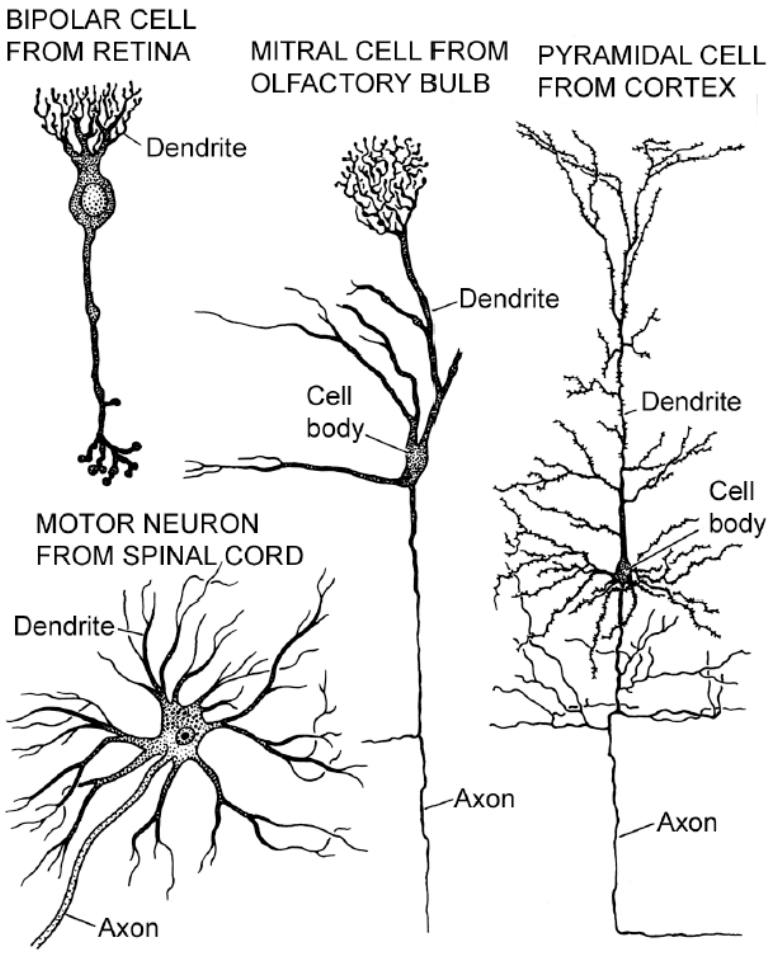
Synapse: 突触

Neurons vs. Other Cells

- Specialized extensions: dendrites and axons
- Neurons communicate with each other through an electrochemical process



Different Neurons

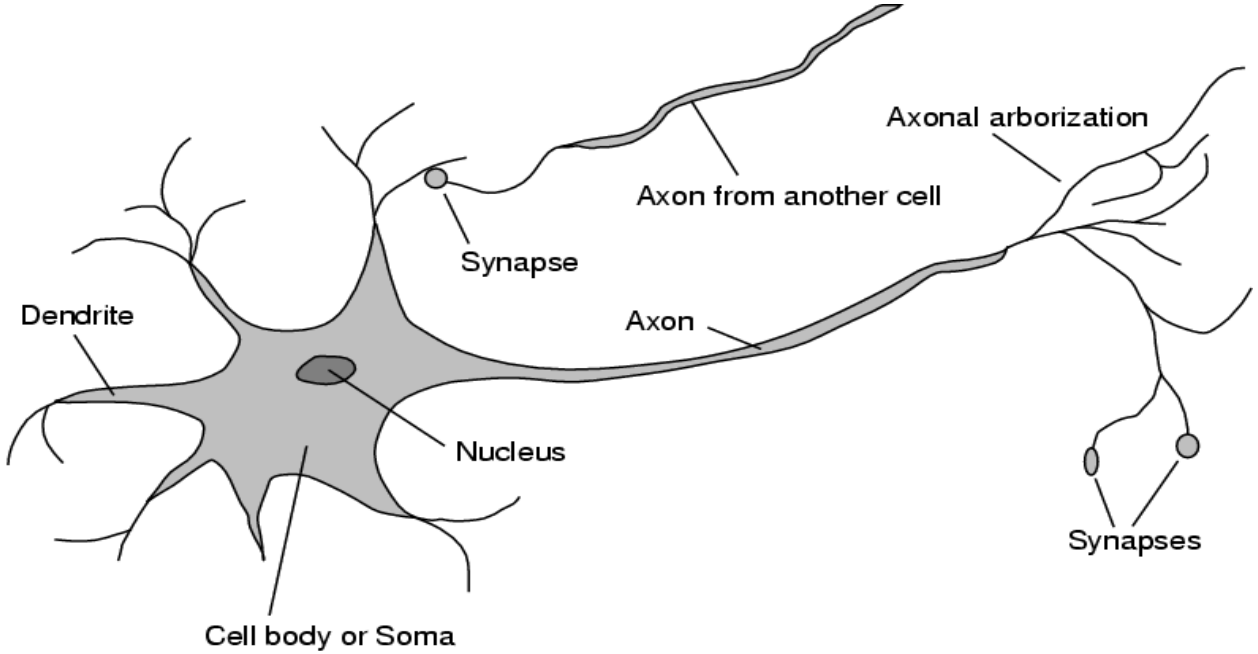


- Neurons come in many different shapes and sizes
- Some of the smallest neurons have cell bodies that are only 4 microns wide, while some of the biggest neurons have cell bodies that are 100 microns wide

Micron=100万分之一米

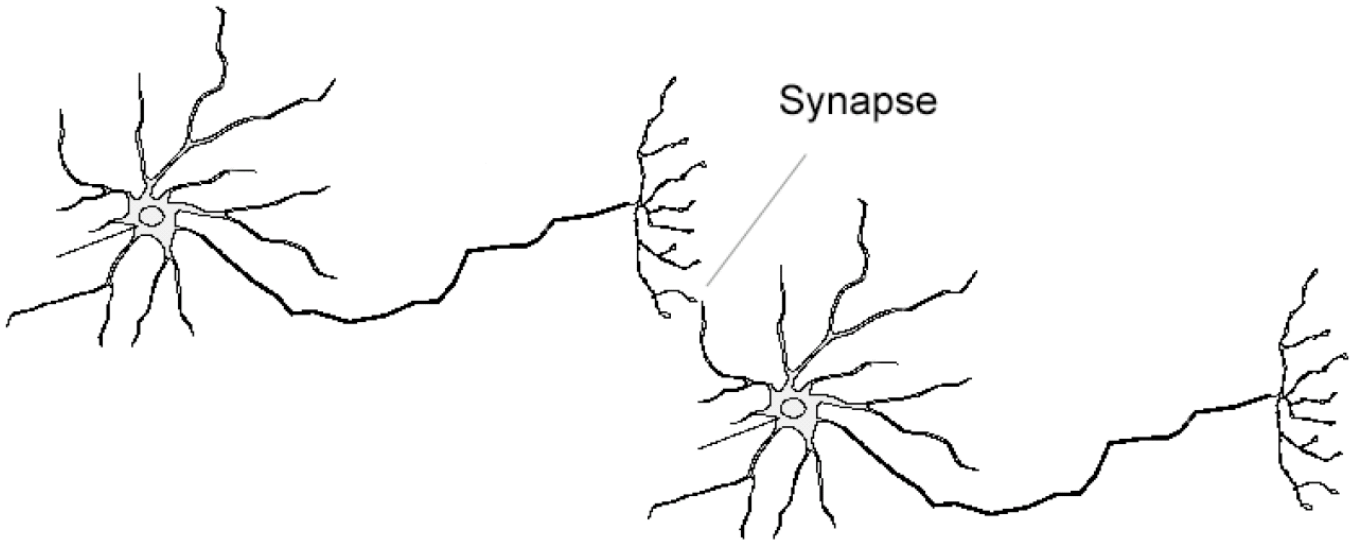
Axons and Dendrites

- Dendrites branch and terminate *in the vicinity of the cell body*, rarely more than a millimeter long and often much shorter
- Axons can extend to *distant targets*, more than a meter away in some instance



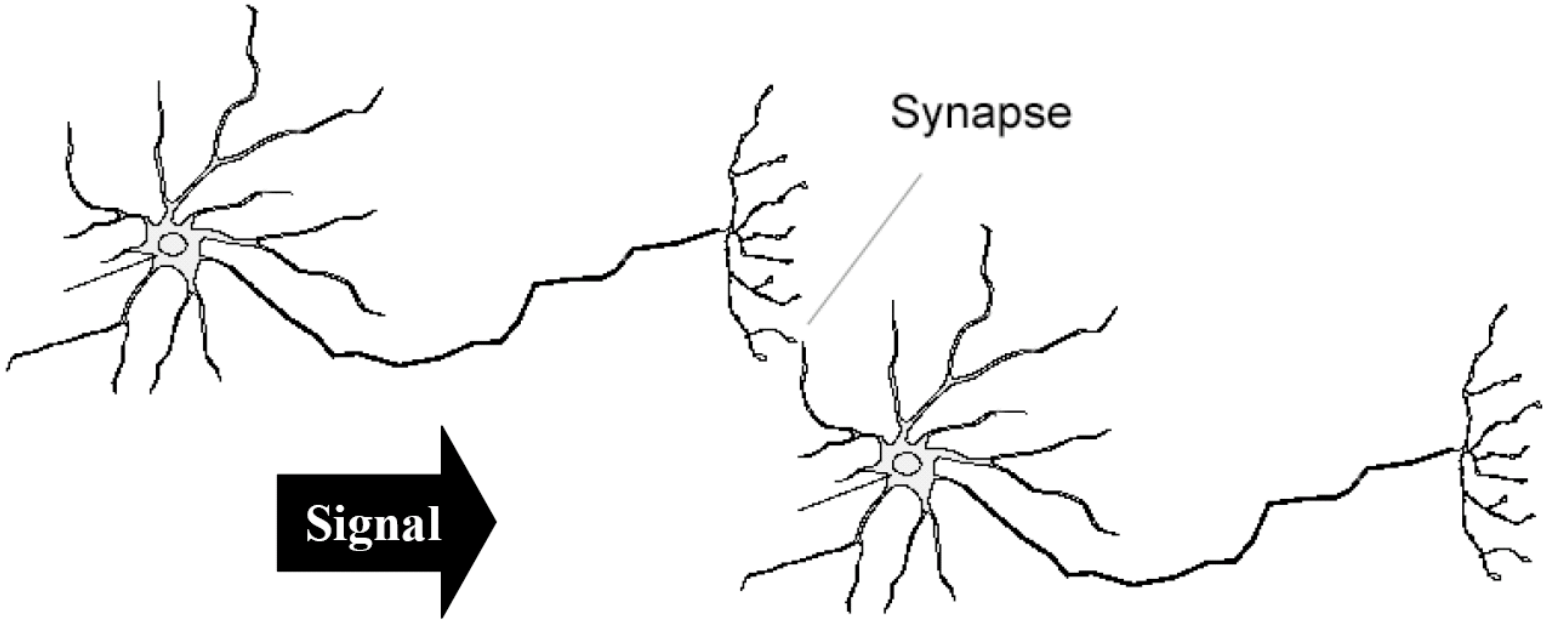
Synapses

- Neurons communicate through specialized junctions called *synapses*

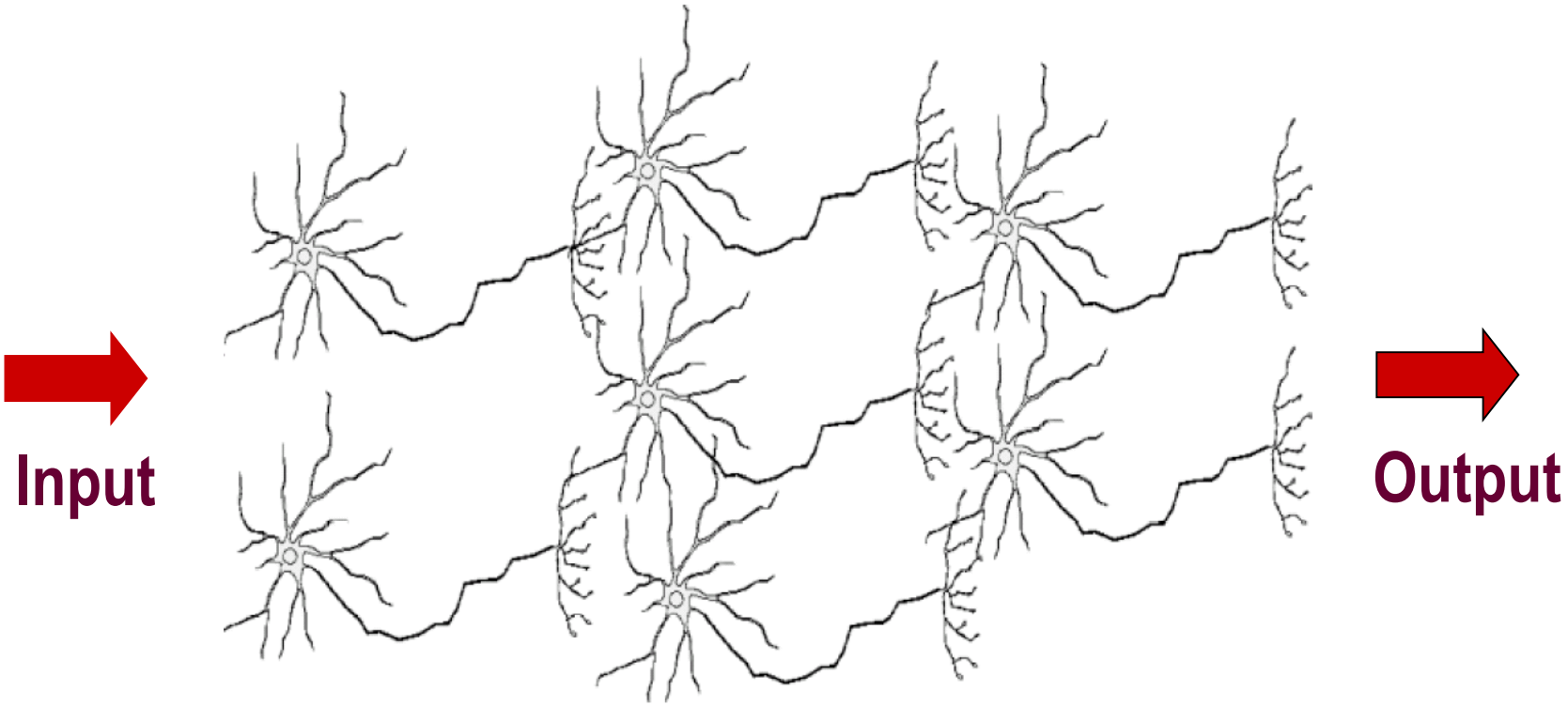


Axonal Signals

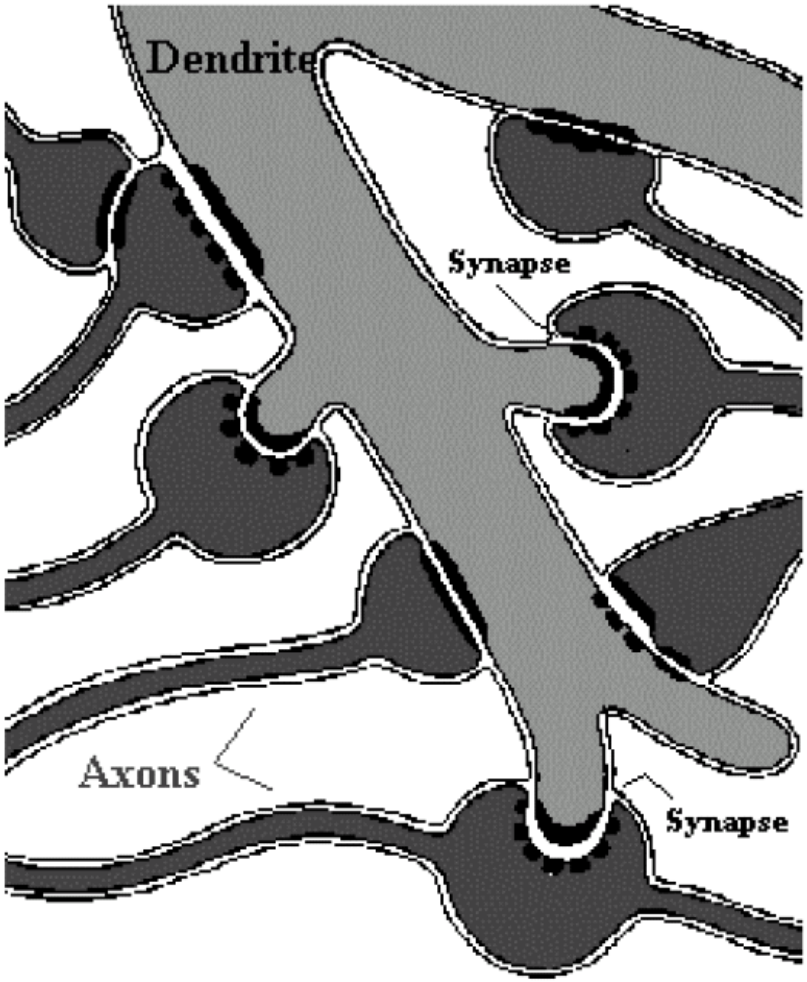
- An electrical signal
- 20 m/s to 100 m/s



A Neural Network

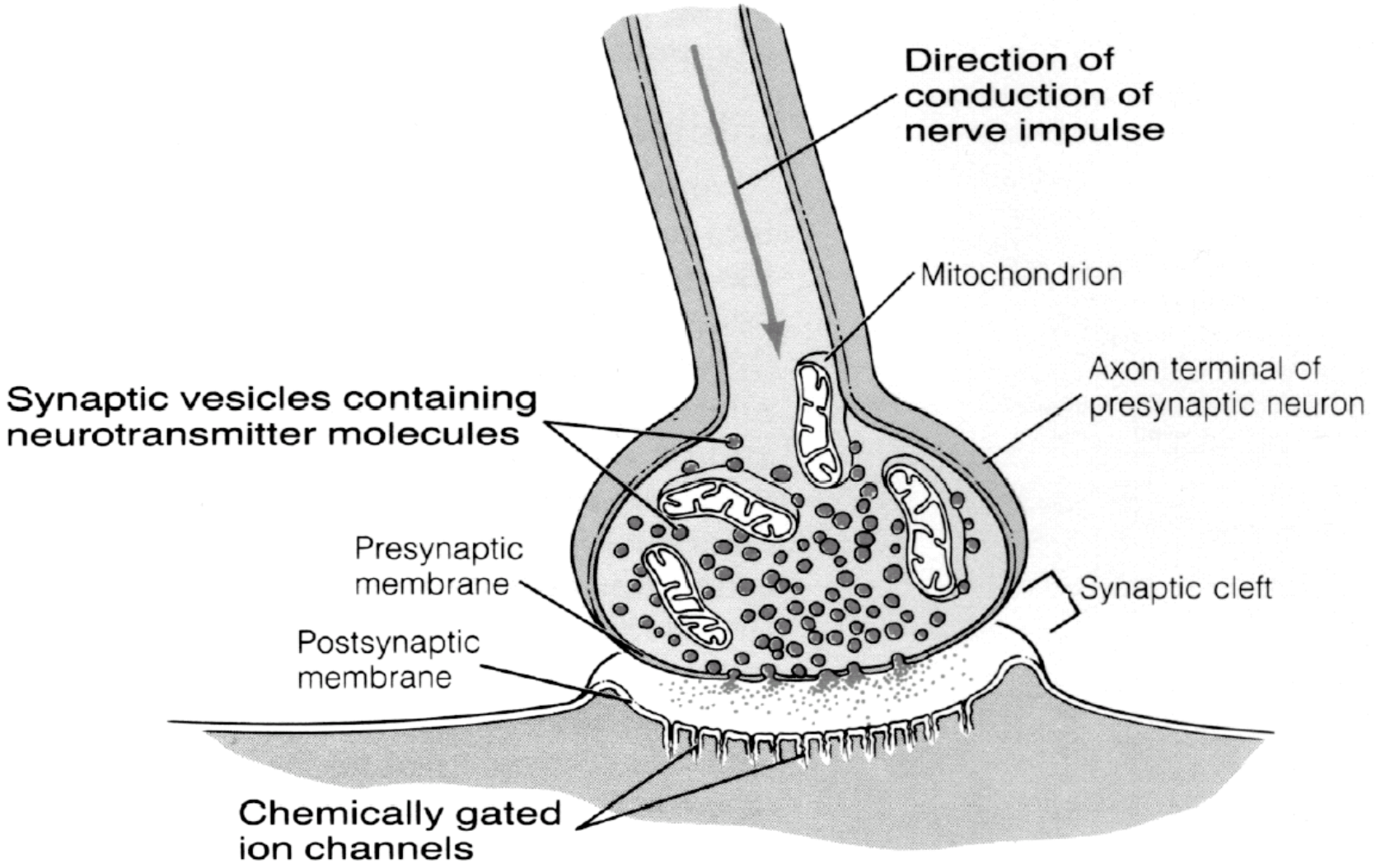


Synapses and Dendrites

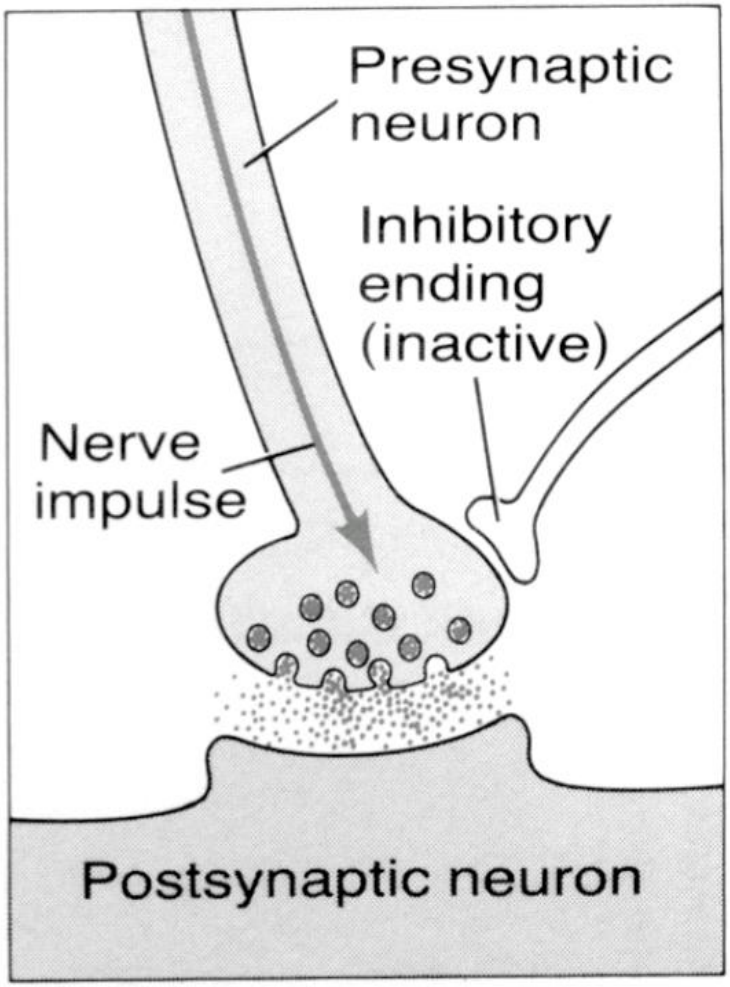


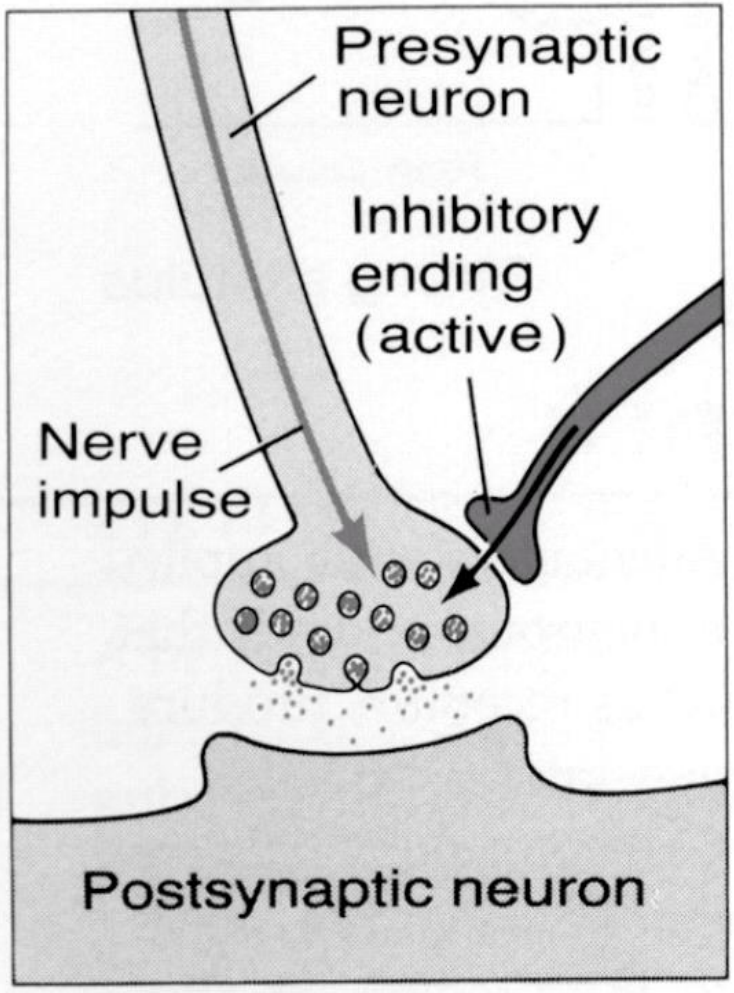
- Axon releases chemical transmitters into the synaptic cleft

Synapse


Direction of conduction of nerve impulse
Mitochondrion
Axon terminal of presynaptic neuron
Synaptic vesicles containing neurotransmitter molecules
Presynaptic membrane
Synaptic cleft
Postsynaptic membrane
Chemically gated ion channels

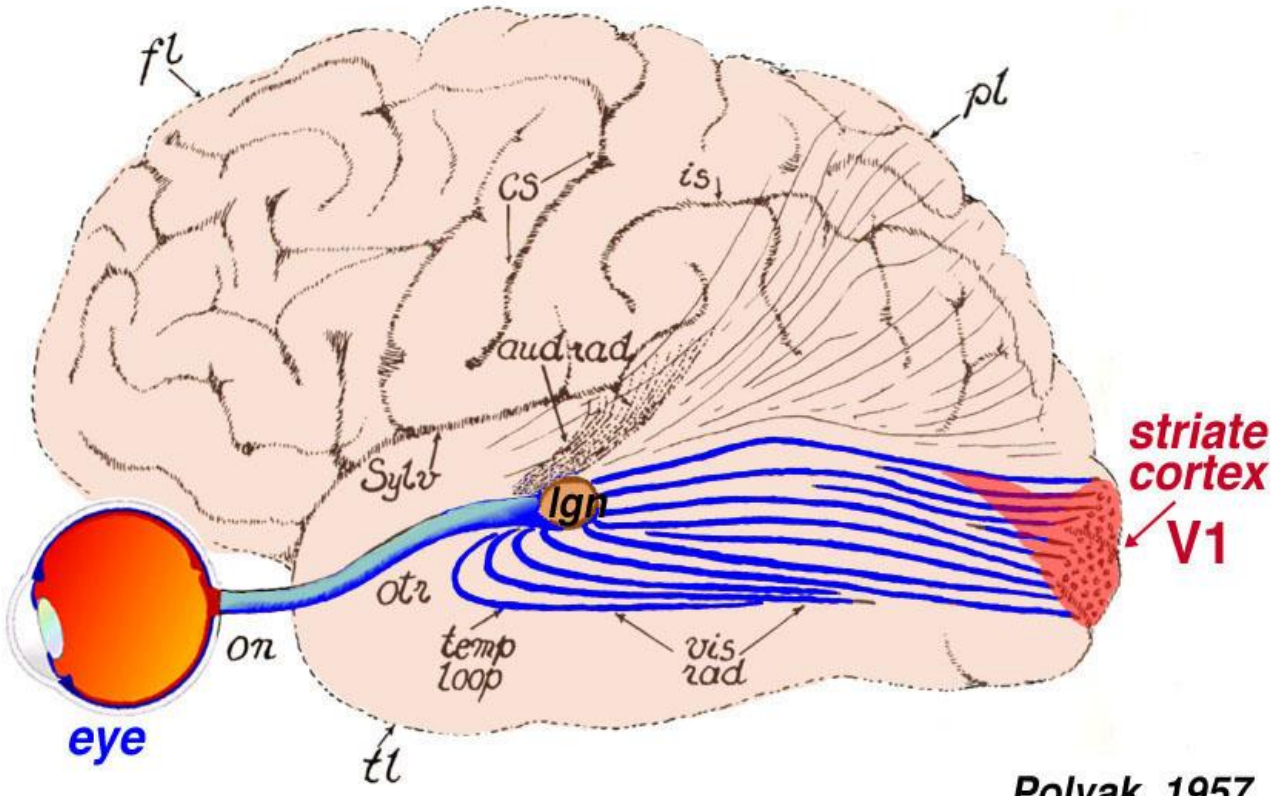
Inhibition


Presynaptic neuron
Inhibitory ending (inactive)
Nerve impulse
Postsynaptic neuron

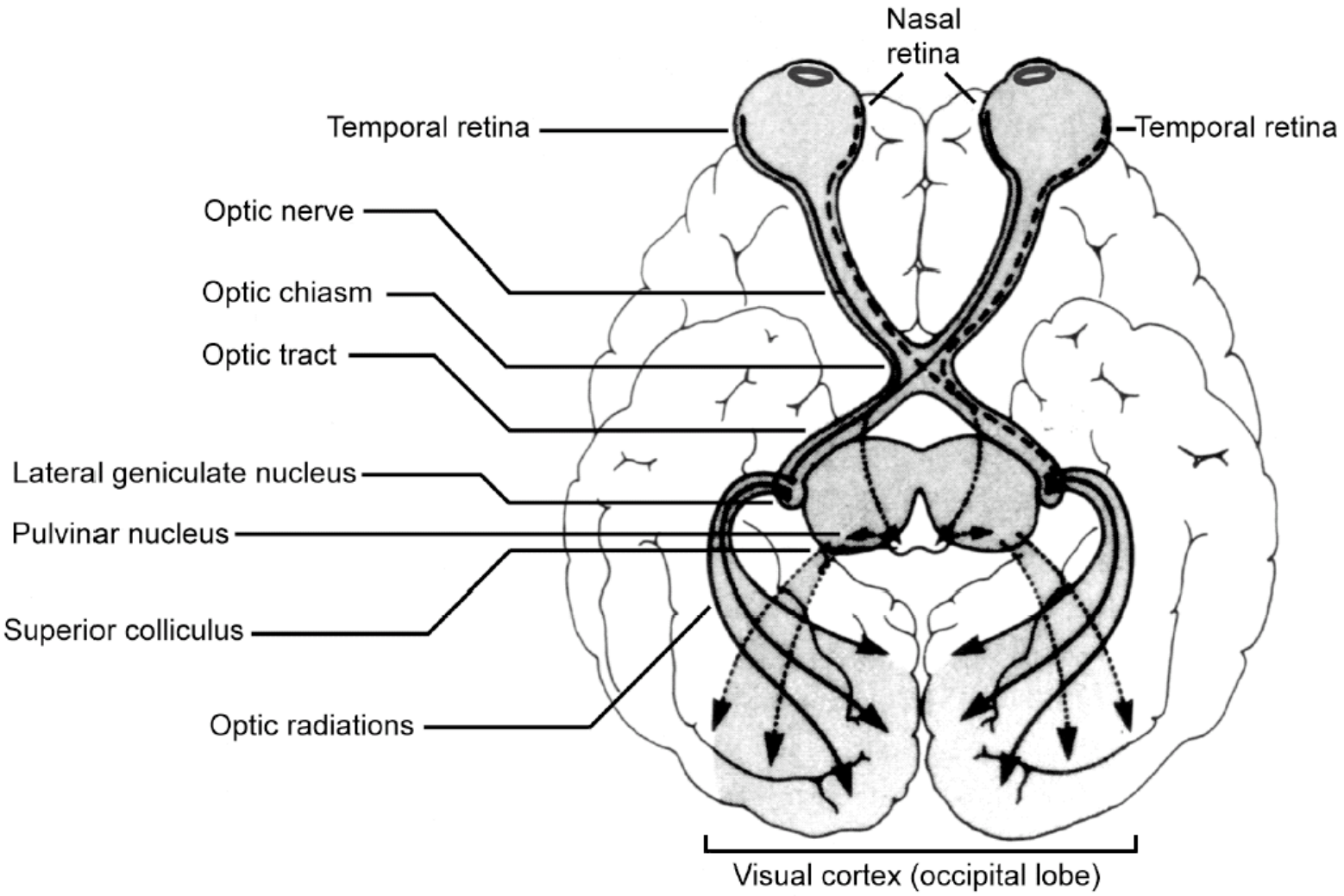

Presynaptic neuron
Inhibitory ending (active)
Nerve impulse
Postsynaptic neuron

视觉皮层（Visual Cortex）

- 至少有70～80%以上的外界信息是由视觉系统接受、处理和感知的；
- 人类大脑皮层约一半的区域参与视觉信息的分析。

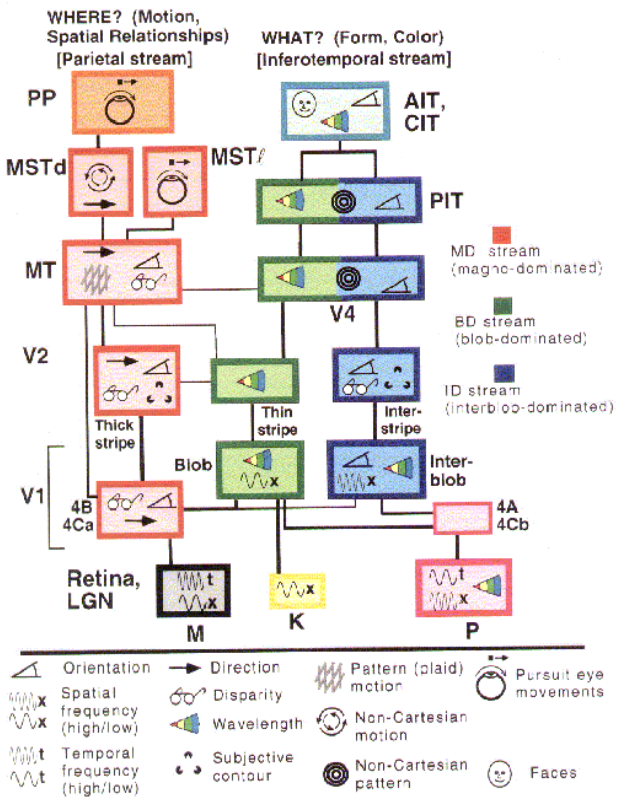


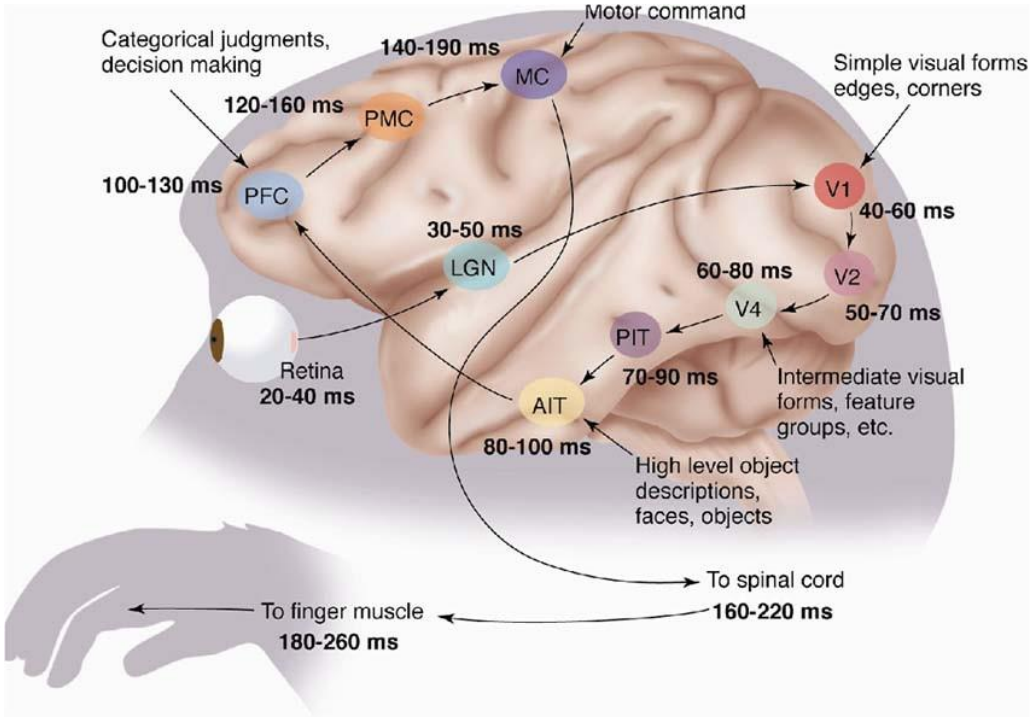
Visual Cortex

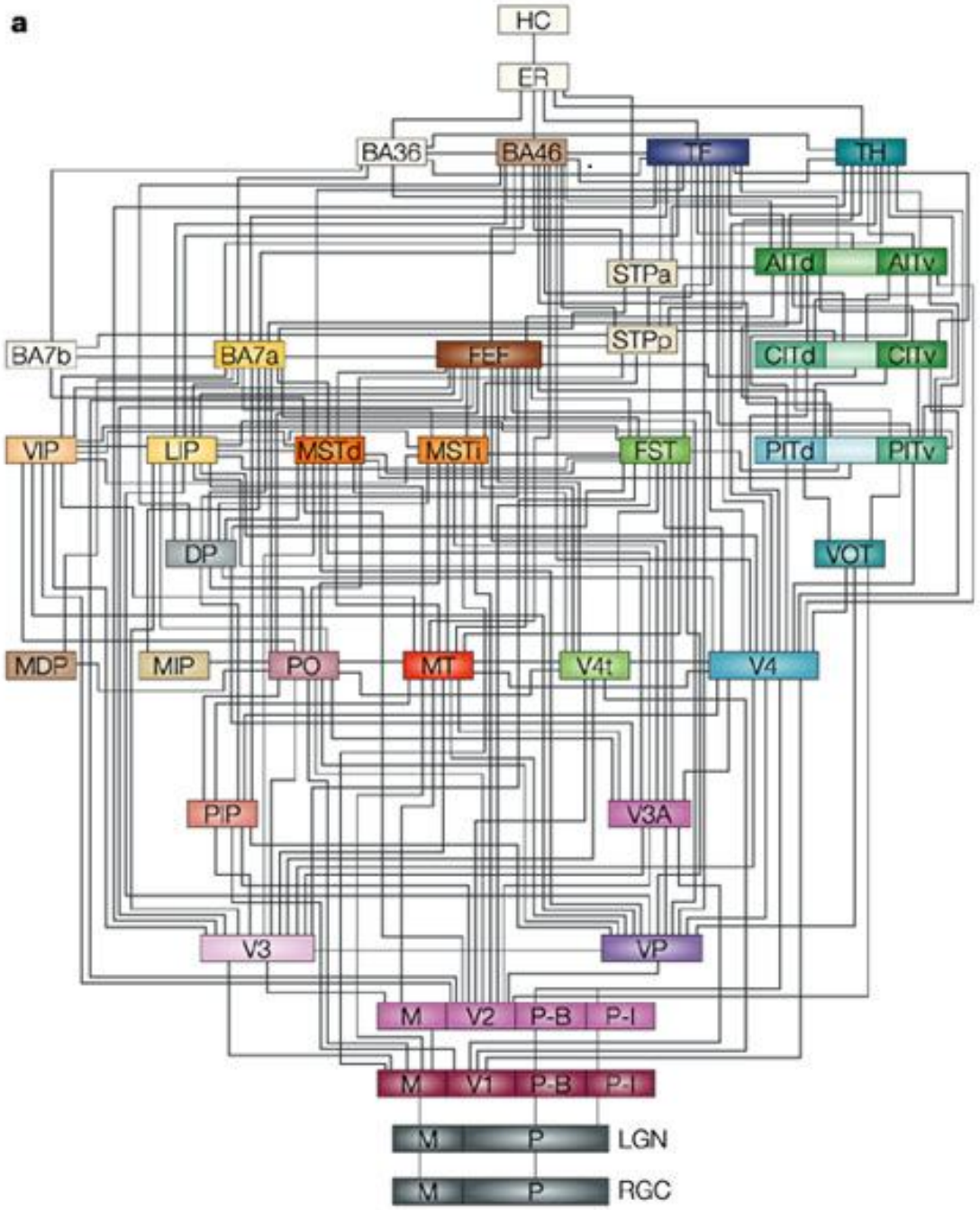


The Mammalian Visual cortex is Hierarchical

- The ventral (recognition) pathway in the visual cortex has multiple stages
- Retina - LGN - V1 - V2 - V4 - PIT - AIT ....
- Lots of intermediate representations

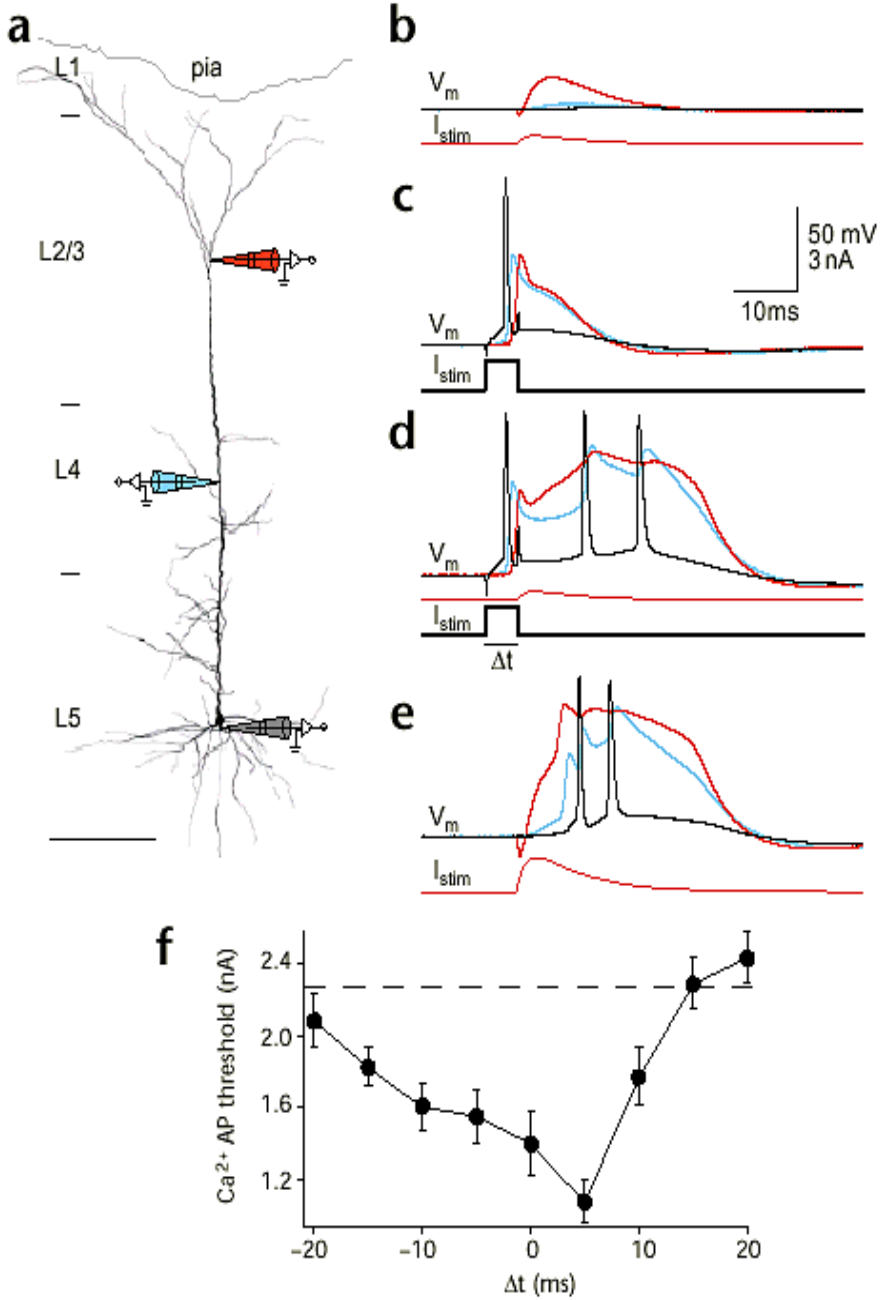


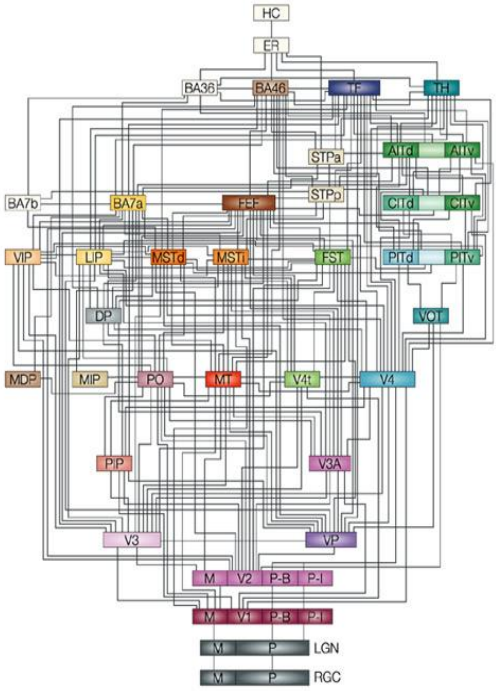


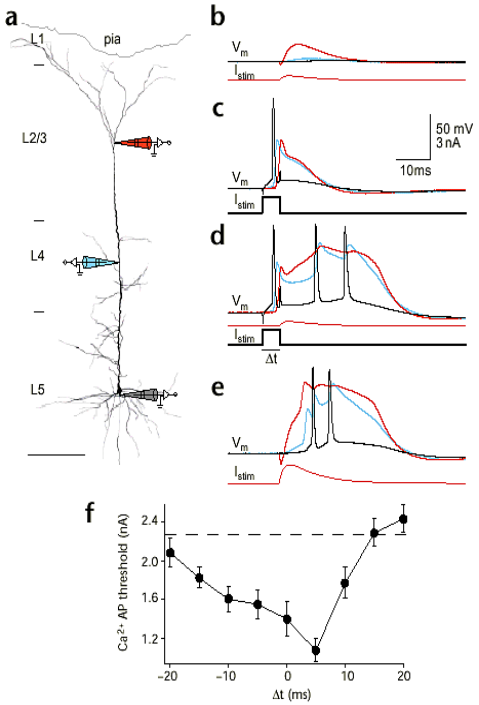


猴视觉皮层既平行又分级的30多个视觉皮层区域间连接的示意图

- 30多个视皮层区
- 10多个等级
- 305条视觉联系


a
L1
pia
L2/3
L4
L5
b
V_m
I_{stim}
c
50 mV
3 nA
10ms
V_m
I_{stim}
d
V_m
I_{stim}
Δt
e
V_m
I_{stim}
f
Ca^{2+} AP threshold (nA)
2.4
2.0
1.6
1.2
−20
−10
0
10
20
Δt (ms)





Our knowledge of organizing neurons in system level is rather poor !

Emergence: From Chaos to Order

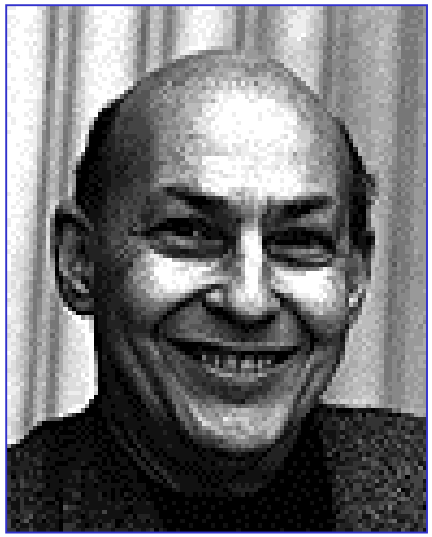
John H. Holland (1998)

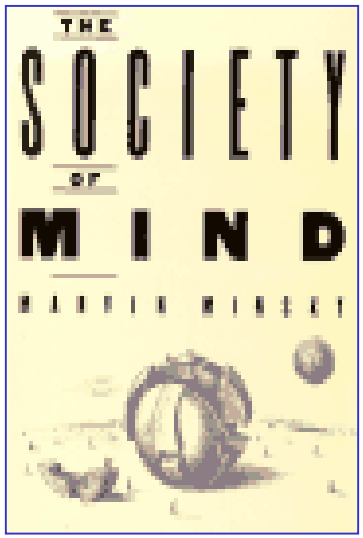
A Theory of Emergence

"We are everywhere confronted with emergence in complex adaptive systems: ant colonies, network of neurons, Internet…, where the behavior of the whole is much more complex than the behavior of the parts."

J. H. Holland, Emergence: From Chaos to Order (1998)

Marvin Minsky (1986)





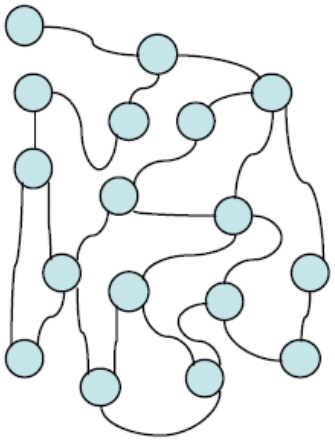
Emergence of Intelligence

"This book tries to explain how minds work. How can intelligence emerge from non-intelligence ? To answer that, we'll show that you can build a mind from many little parts, each mindless by itself"

Ant Colony

What is a Neural Network ?

- A neural network is a network of many simple processors (neurons, units)
  - The units are connected by *connections*
  - Each connection has a number weight associated with it
  - The units operate only locally – on their weights and the inputs they receive through the connection



- Connectionism
- Parallel distributed processing
- Neural computation
- Neurocomputing

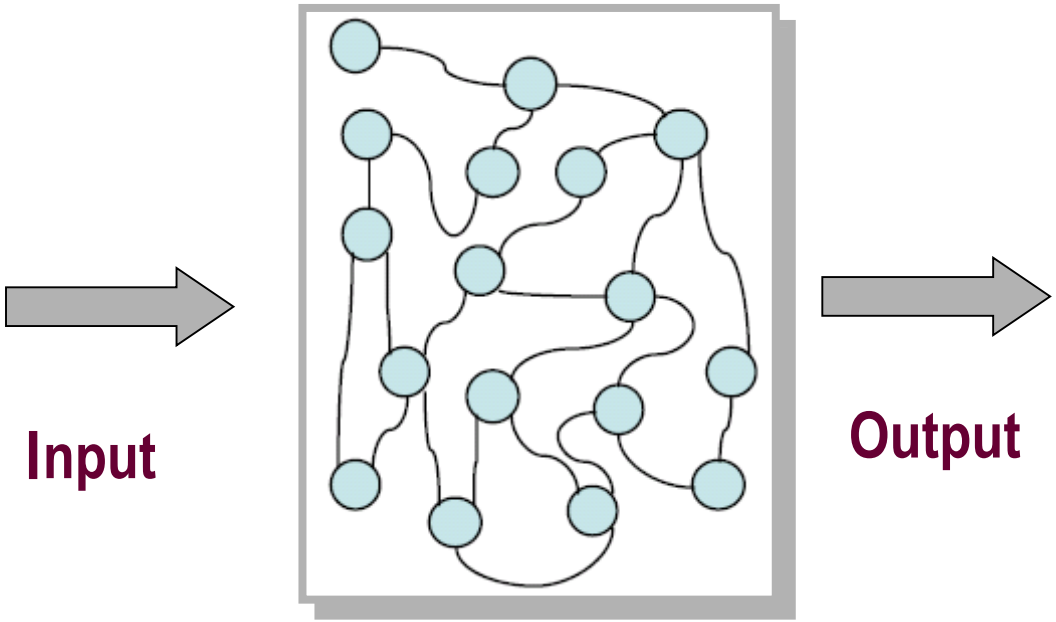
What is a Neural Network? (Cont.)

- A NN is a massively parallel distributed processor made up of simple processing unit, which has a natural propensity for storing experimental knowledge and making it available for use. It resembles the brain in two respects:
  - Knowledge is acquired by the network from its environment through *a learning process*
  - Interneuron connection strengths, known as *synaptic weights*, are used to store the acquired knowledge

What is a Neural Network ? (Cont.)

- Some NNs are models of biological NNs and some are not
- The inspiration for the field of NNs came from the desire to produce artificial systems capable of sophisticated computations similar to those that the human brain routinely performs, and thereby also to enhance our understanding of the brain

Information Processing



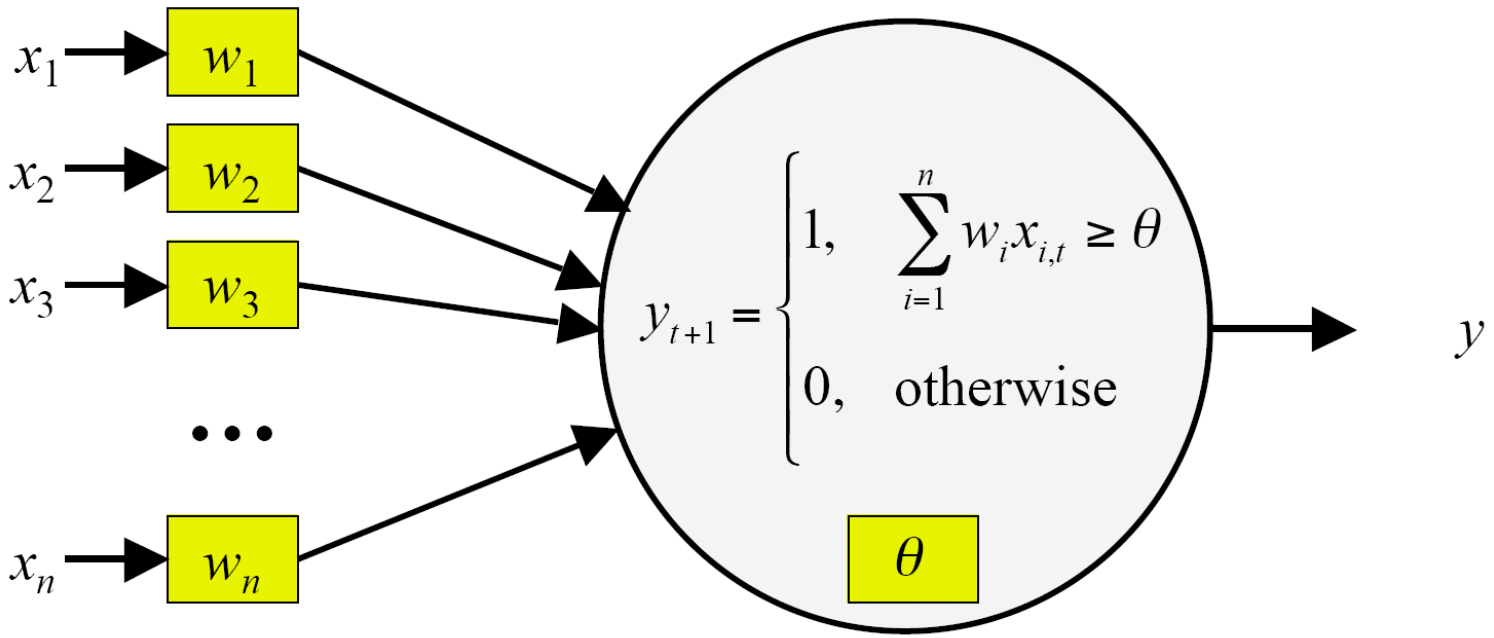
Neural Network History

- 1943: McCulloch and Pitts Model
- 1949: Hebb's rule
- 1962: Rosenblatt's learning Perceptron
- 1969: Minsky and Papert published their book "Perceptrons" on the limitations of perceptrons
- 1970s: Kohonen, Anderson, Grossberg, Amari, Fukushima
- 1980s: Rumelhart and McClelland, Hopfield, etc.
- 1990s: NNs have found many practical applications
- 2006: Hinton, LeCun, Bengio, etc. deep learning

Professors Amari and Fukushima in ICONIP2011

Neural Network History -1

- 1943: McCulloch and Pitts develop basic models of neurons
  - The first artificial neuron



Neural Network History -2

1943: Warren McCulloch and Walter Pitts

Networks of logical threshold units (all or nothing responses) can perform logic calculations. Any finite logical expression can be realized by these McCulloch-Pitts neurons.

Describes a true connectionist model, with simple computing elements, arranged largely in parallel, doing powerful computations with appropriately constructed connections.

Neural Network History -3

1949: Donald O. Hebb

The organization of behavior was the first explicit statement of a physiological learning rule for synaptic modification (since become known as the Hebb rule).

"When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased."

Neural Network History -4

1950-1970: Many papers on associative memory models
(Taylor, Willshaw, Longuet-Higgins, Anderson, Kohonen, Nakano).

Correlation matrix memories

1956: F. Rosenblatt

The perceptron model and the perceptron convergence algorithm

Described a learning machine with simple computing elements that was potentially capable of complex adaptive behaviors.

1960: Widrow and Hoff

Introduced the least mean square error algorithm (gradient descent) which is the basis for most modern "error correction rules".

Neural Network History -5

1969: Minsky and Papert

Used elegant mathematics to demonstrate that there are fundamental limits on what a one-layer perceptron can compute.

"In the popular history of neural networks, first came the classical period of the perceptron, when it seemed as if neural networks could do anything. A hundred algorithms bloomed, a hundred schools of learning machines contended. Then cam the onset of the dark ages, where, suddenly, research on neural networks was unlived, unwanted, and most important, unfunded. "

Neural Network History -6

1973; 1976: Christoph van der Malsburg

Demonstrated self-organization in computer simulations motivated by topologically ordered maps in the brain.

1980s: Stephen Grossberg

Adaptive resonance theory

1982: John Hopfield

Used the idea of an energy function to formulate a new way of understanding the computation performed by recurrent networks with symmetric synaptic connections. He established the relation between such recurrent networks and an Ising Model used in statistical physics.

1982: Teuvo Kohonen

Self-organizing maps

Neural Network History -7

1983: Sutton, Barto and Anderson

Introduced reinforcement learning and showed that a reinforcement learning system could learn to balance a broomstick in the absence of a helpful teacher.

1986: Rumelhart, Hinton and Williams

Developed the backpropagation algorithm which solved the credit assuagement problem for multi-layer networks, which emerged as the most popular algorithm for the training of neural networks. It was discovered independently also by Parker and LeCun.

Neural Network History -8

2006: Hinton, LeCun, and Bengio

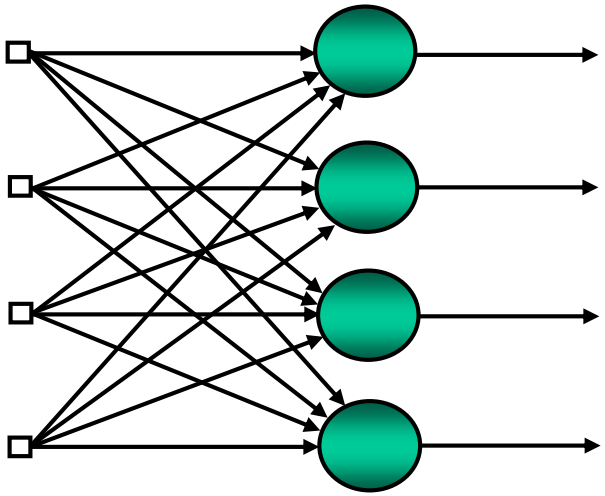
Developed deep learning algorithms and successfully applied to unsupervised feature learner.

This Lesson: Part 2

- Network architectures
- Basic learning rules
- Learning paradigms
- Learning tasks

Network Architectures

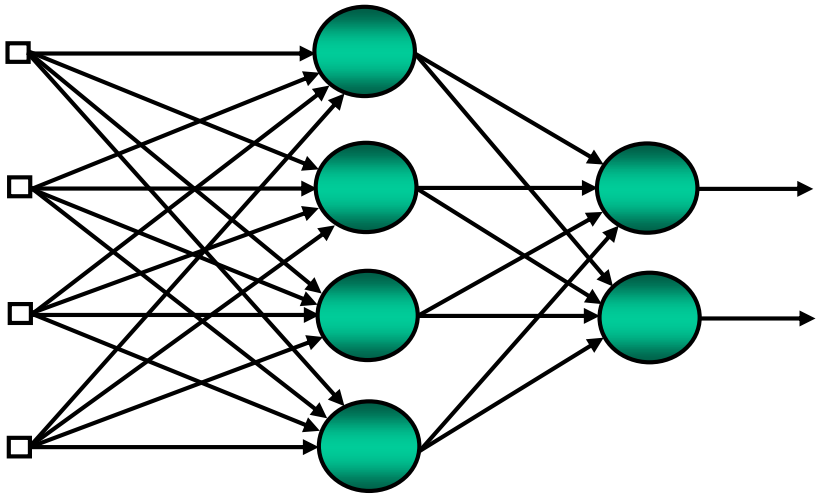
- Single-Layer Feedforward Networks
- Multilayer Feedforward Networks
- Recurrent Networks



Input layer

Output layer

Multilayer Feedforward Networks

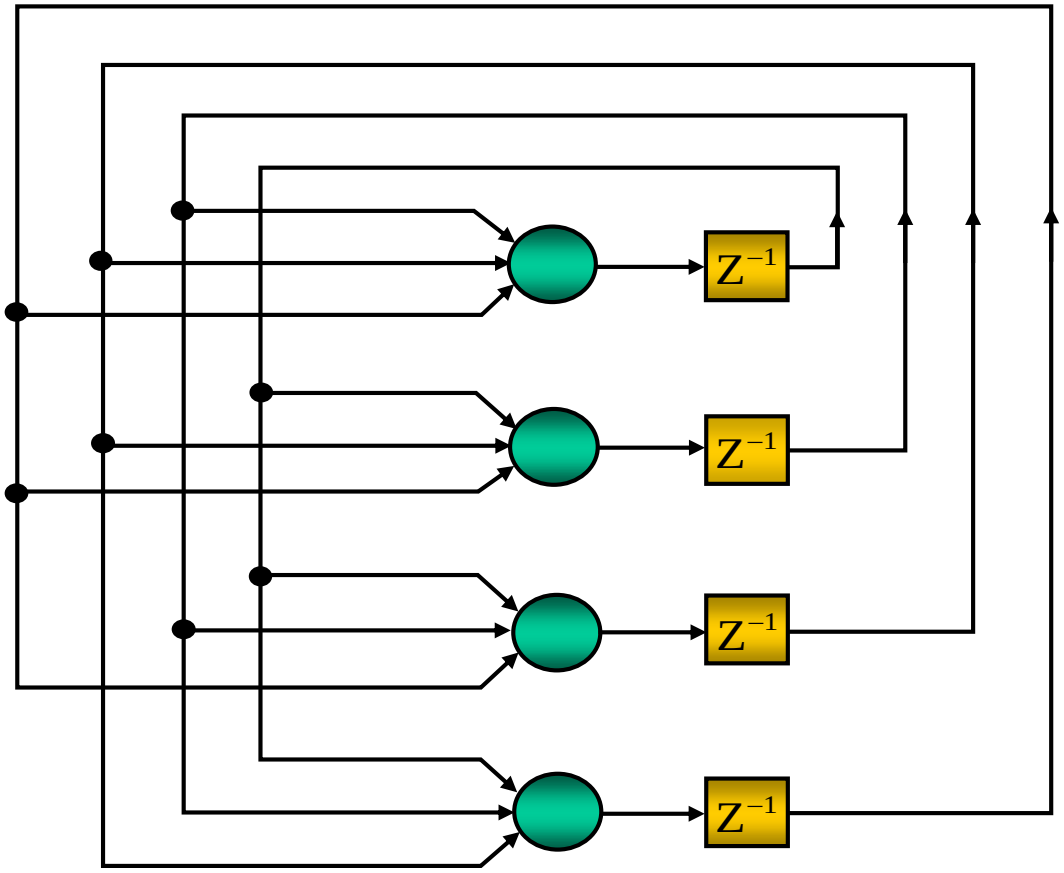


Input layer

Hidden layer

Output layer

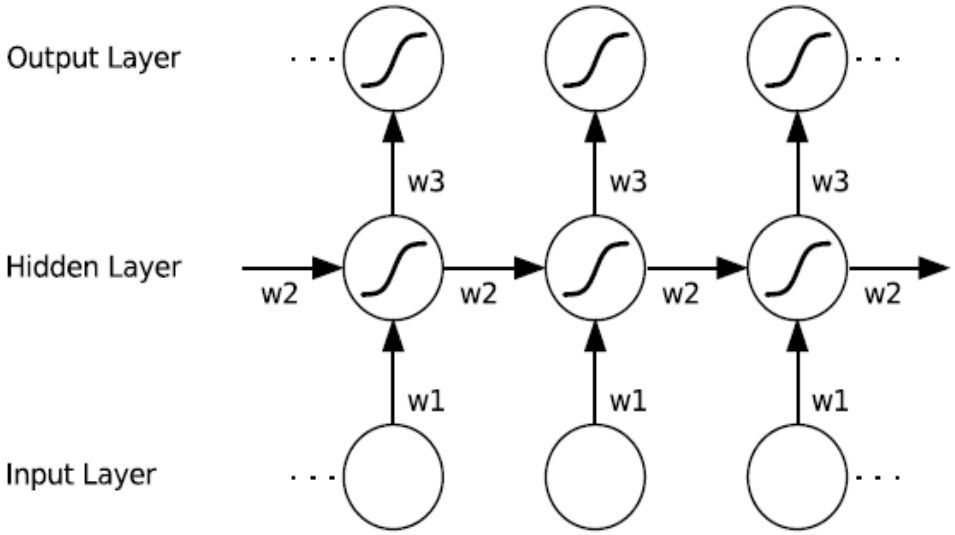
Recurrent Networks

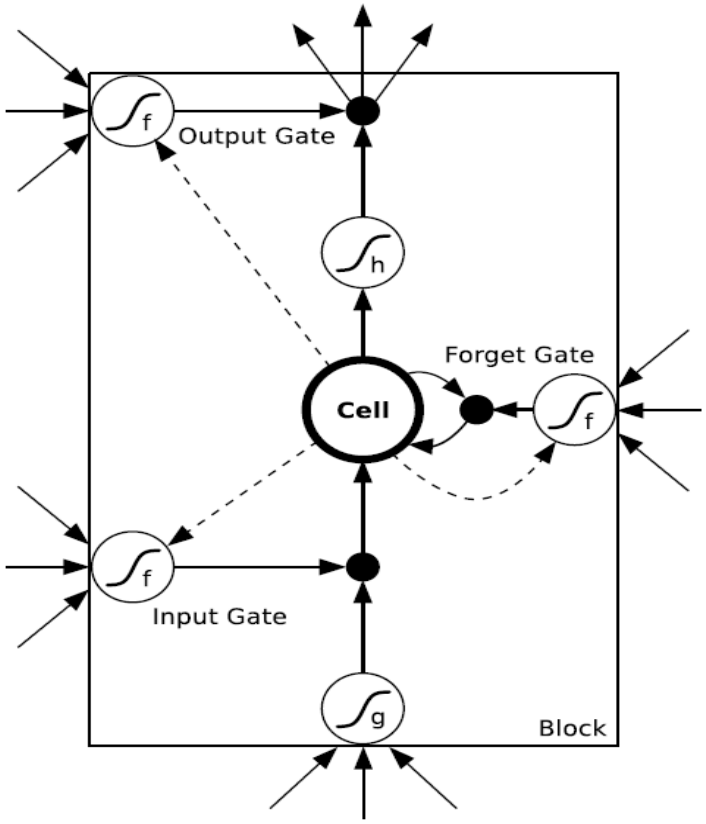


Neurons Unit-delay Operators

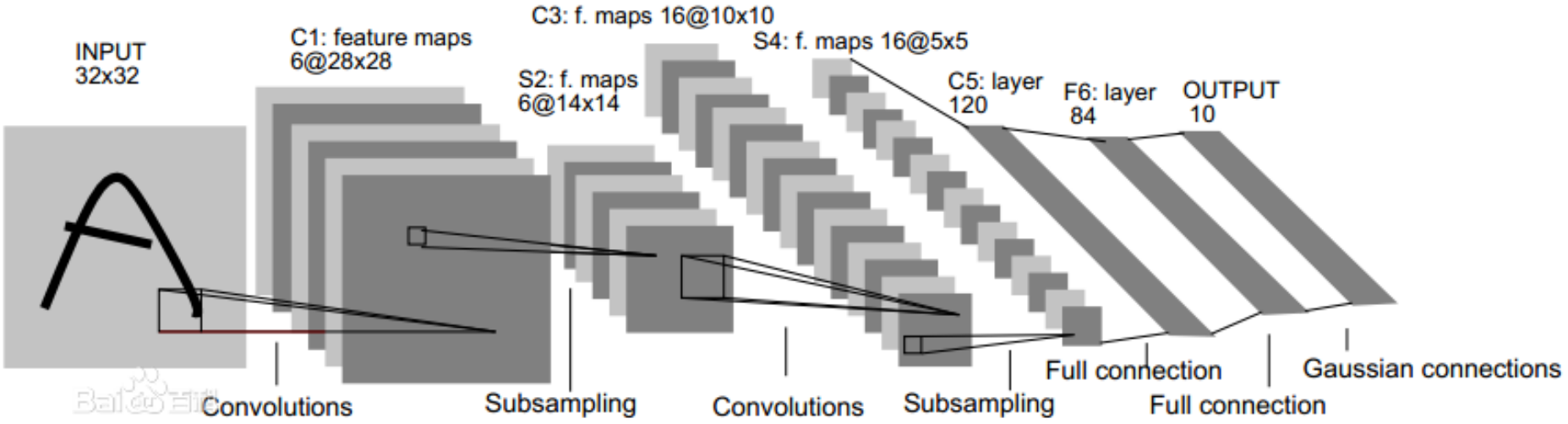
LSTM Neural Network

- Long Short Time Memory (LSTM) is one kind of recurrent neural network which can store states over a long period of time.





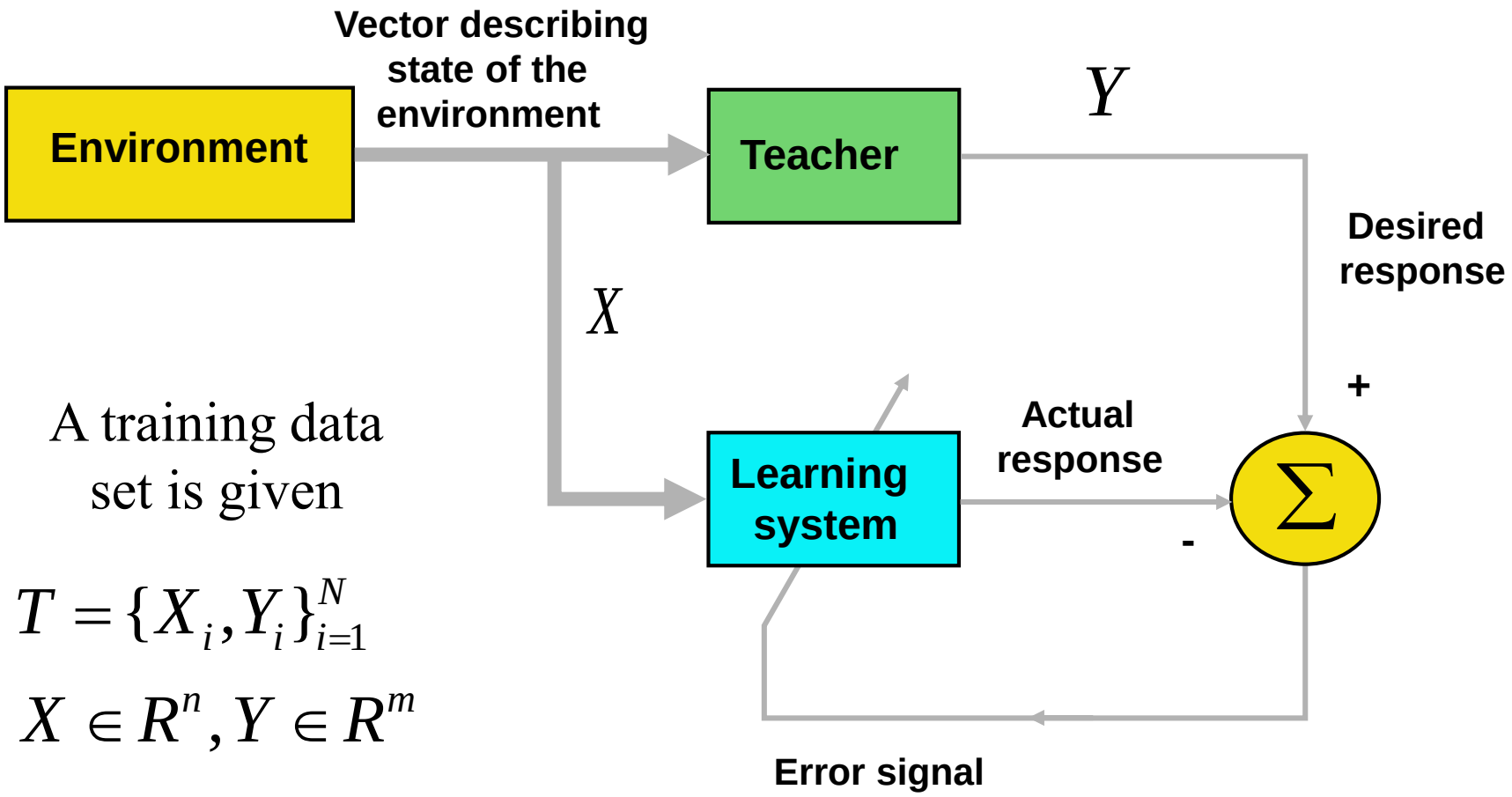
Deep Learning Structure



Learning Paradigms

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

Supervised Learning



A training data set is given

$$T = \{X_i, Y_i\}_{i=1}^{N}$$

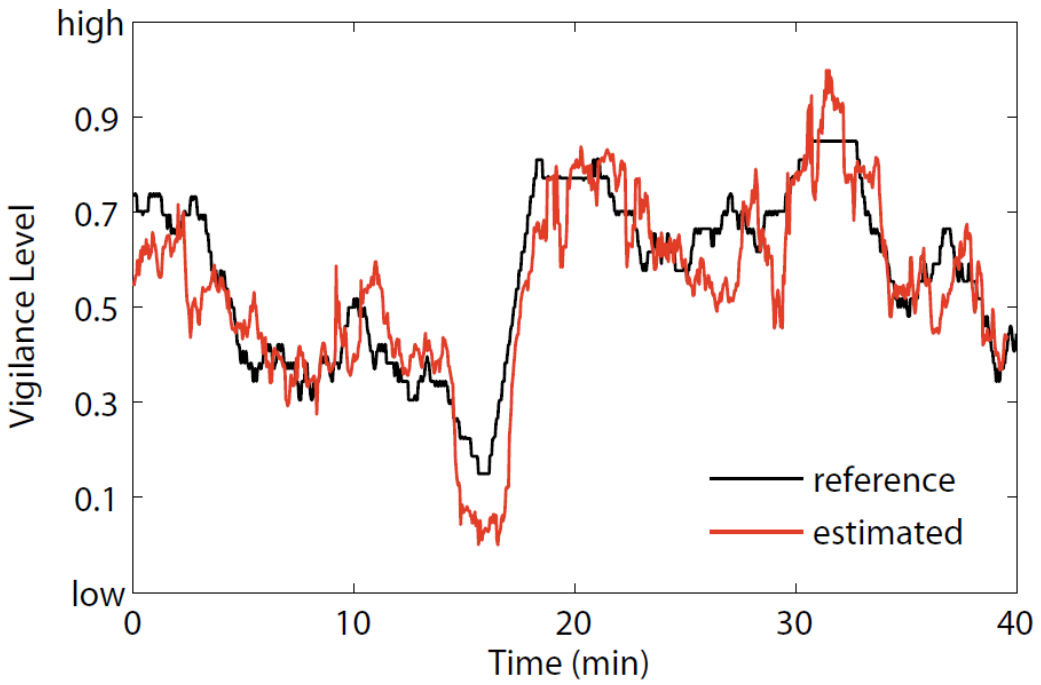
$$X \in R^n, Y \in R^m$$

监督学习的两类学习任务

- 分类(Classification): 教师信号是离散值
- 预测 (Prediction): 教师信号是连续值

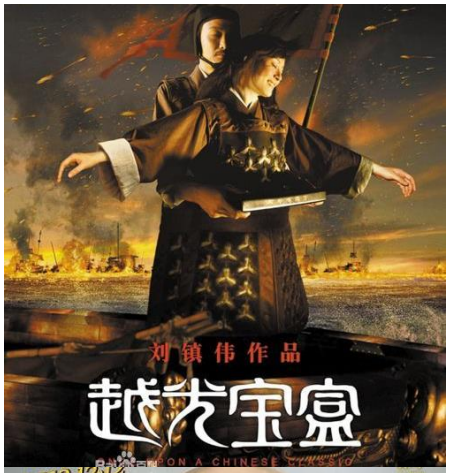
  也叫回归 (Regression)，或

  函数逼近 (Function Approximation)

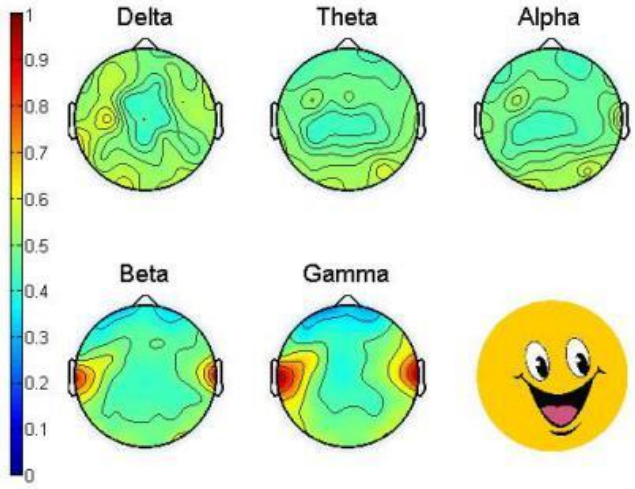


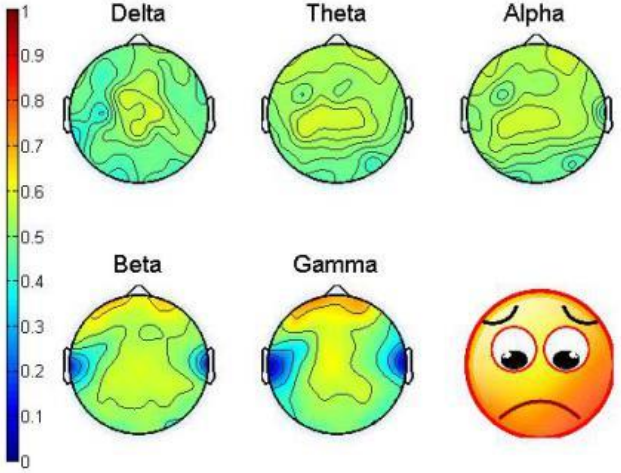
实际的分类问题：基于脑电的情绪识别

Positive movies



Negative movies

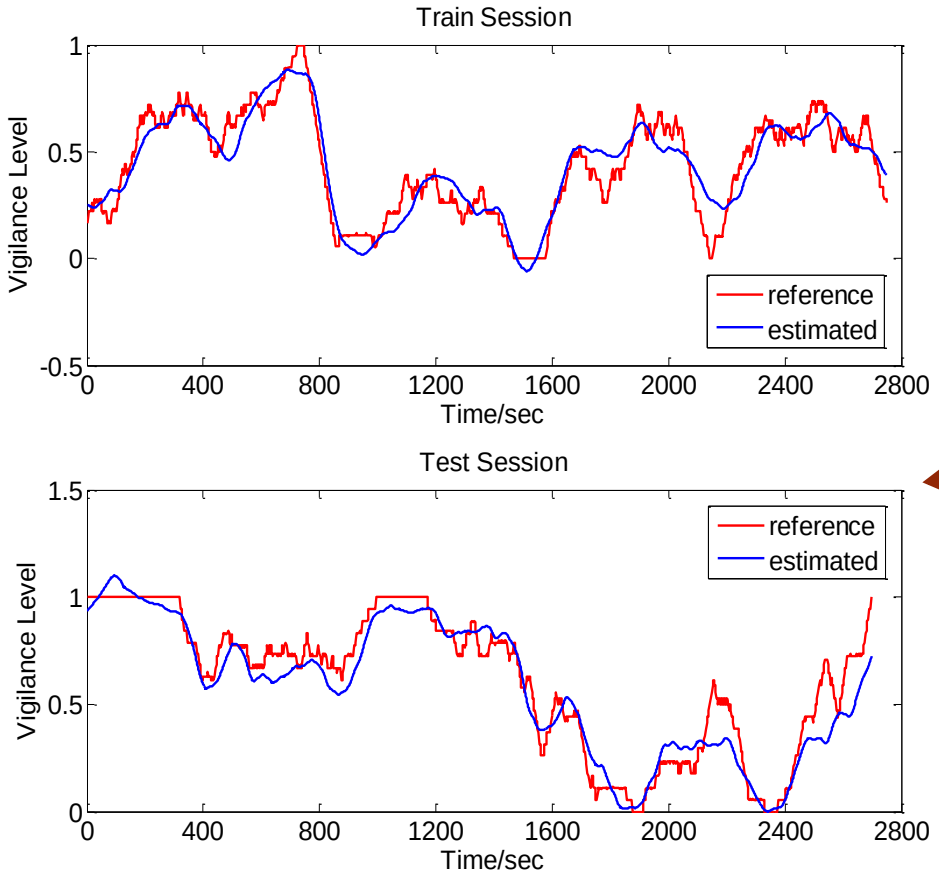




实际的预测问题：司机疲劳检测

实际的预测问题：司机疲劳检测

- Kernel-based supervised regression method



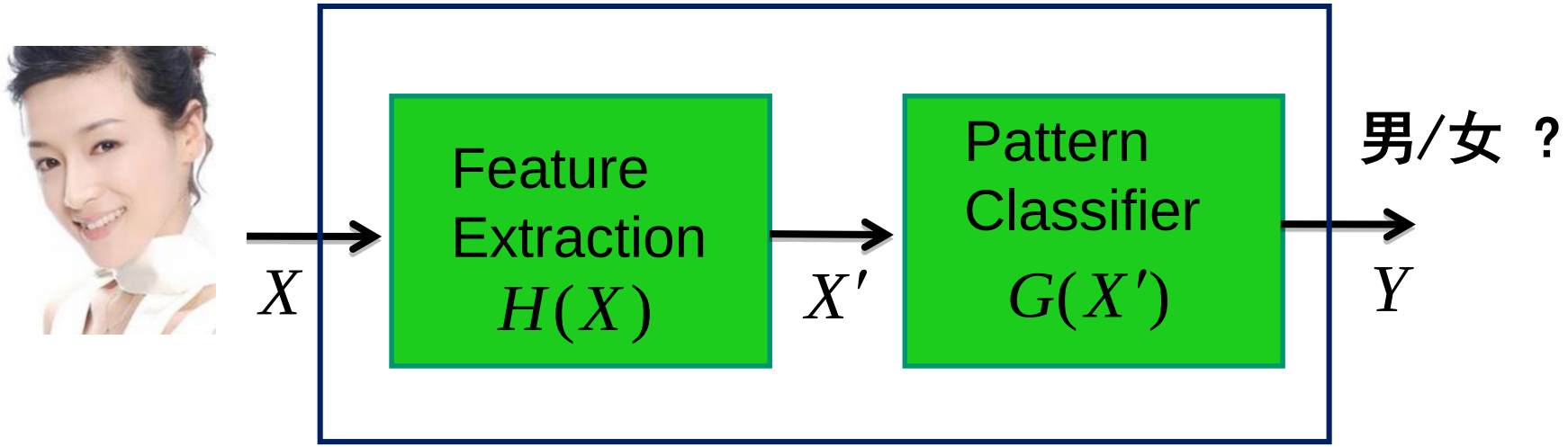
Data Set	Training	Test
1	0.8538	0.7773
2	0.9207	0.8213
3	0.9357	0.8973
4	0.9306	0.8945
5	0.9464	0.9560
6	0.9376	0.9227
7	0.9414	0.8569
8	0.8245	0.8282
9	0.8451	0.8021
Mean	0.9040	0.8618

Training session and test session are two different sessions

实际的分类问题：性别识别与年龄估计

基于机器学习的性别识别系统

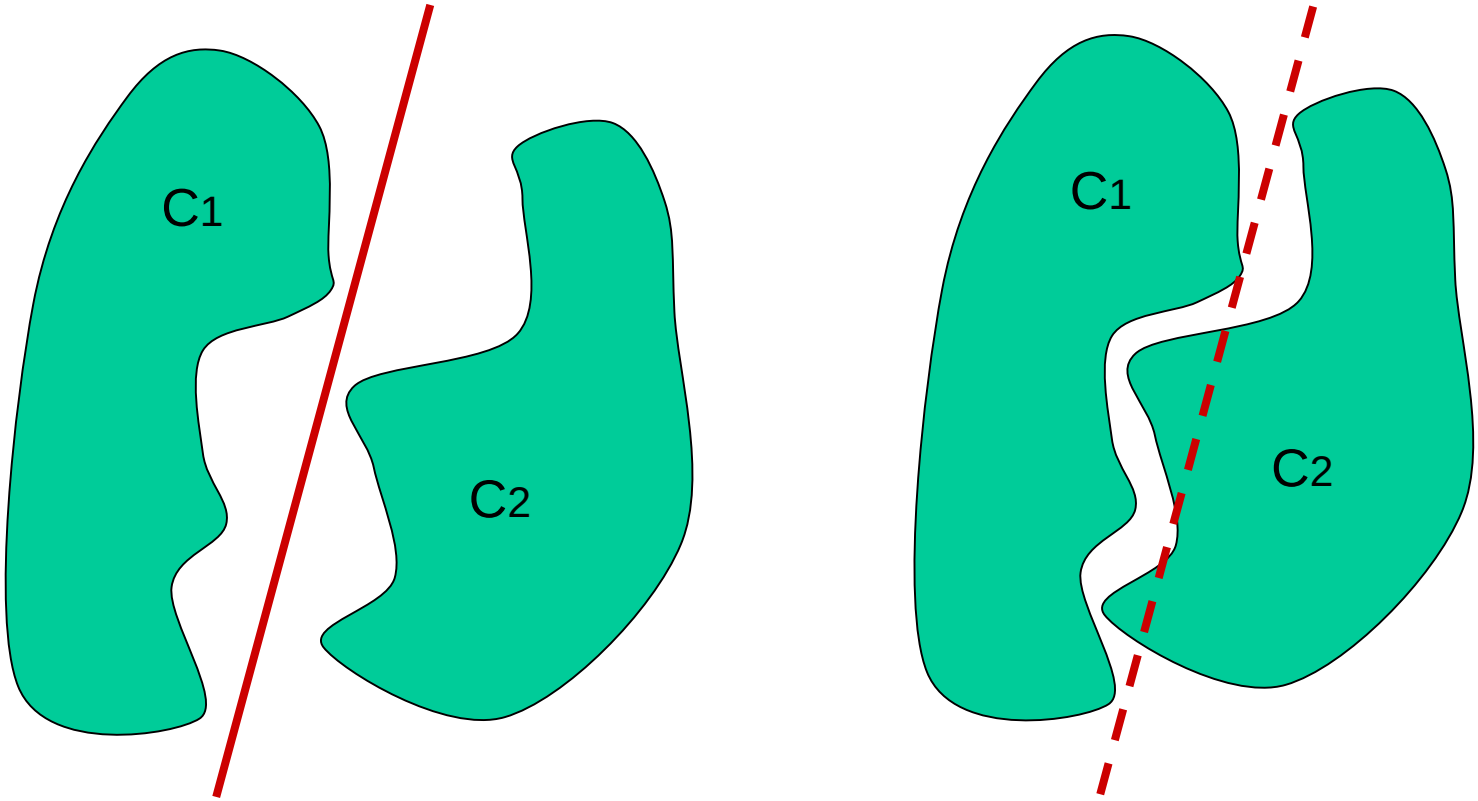
- 系统 输入：人脸图像（X）
- 系统输出：男/女（Z）
- 问题：如何用机器学习方法构造系统？



性别识别系统

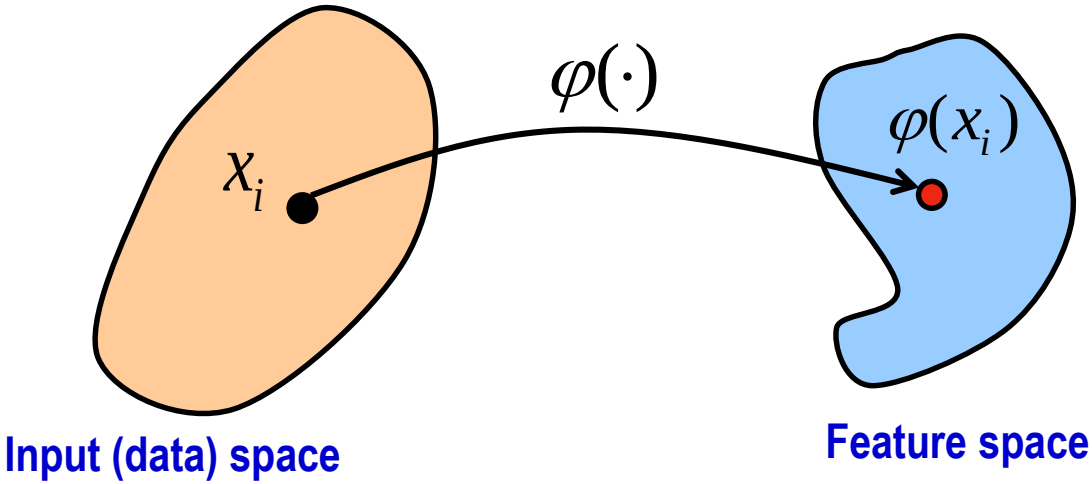
线性可分与非线性可分

- 线性可分(Linear Separable)
- 线性不可分（Linear Non-separable）

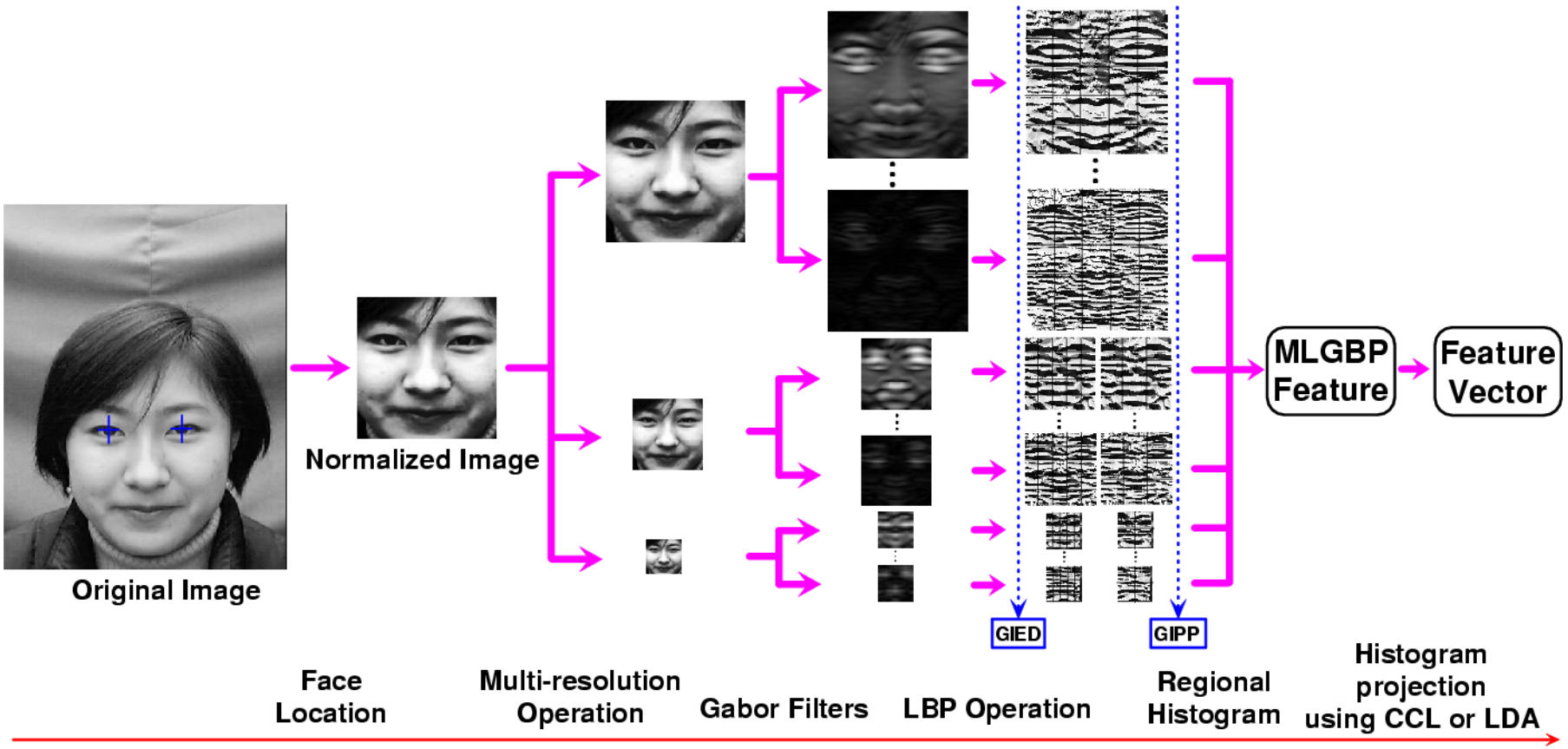


特征提取与特征选择

- 特征提取或抽取（Feature Extraction）
  - 对某一模式的一组测量值进行变换或映射以突出该模式具有代表性特征的方法
- 特征选择（Feature Selection）



人脸特征（ MLGBP）


Original Image
Normalized Image
GIED
GIPP
MLGBP Feature
Feature Vector
Face Location
Multi-resolution Operation
Gabor Filters
LBP Operation
Regional Histogram
Histogram projection using CCL or LDA

Neural Network Journals

- IEEE Trans. Neural Networks and Learning Systems
- Neural Computation
- Neural Networks
- Neurocomputing

- IEEE Trans. Pattern Analysis and Machine Intelligence
- Machine Learning
- Journal of Machine Learning Research

Neural Network Conferences

- Neural Information Processing Systems (NIPS)
- IEEE/INNS International Joint Conference on Neural Networks (IJCNN)
- IEEE World Congress on Computational Intelligence （WCCI）
- International Conference on Neural Information Processing (ICONIP)

Deep Learning Frameworks

- With the rapid developments of the deep learning, different deep learning frameworks are developed to make it easy to build deep learning solutions.

  - Tensorflow
  - Keras
  - Caffe
  - Torch
  - PyTorch
  - MXNet
  - ...

Instructor and Slides

- My Ph.D. student：Wei Liu（刘伟），Yun Luo（罗赟）
- Email：liuwei@liujr.com；angeleader2012@163.com
- Send email to Wei Liu:
  - Your name
  - Your student ID
  - Your email
- You can download the slides from the following FTP:

  ftp://bcmi.sjtu.edu.cn:2122/

  User name: nnet

  Password: nnet