Understanding Deep Neural Networks

Section One: Introduction

Zhang Yi, *IEEE Fellow*
Autumn, 2018

Outline

- Concepts
- An example: handwritten digits recognition
- How does a child recognize the handwritten digits
- Introduction to brain structure
- Discussions

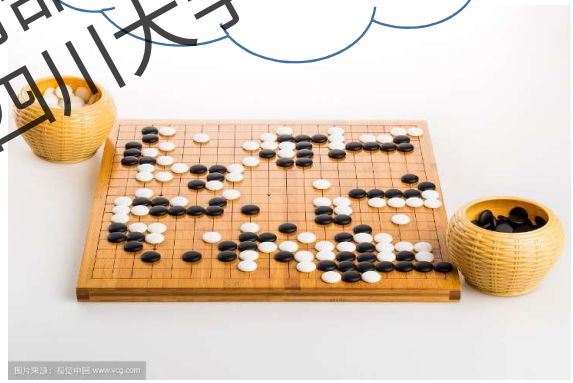
Concepts

March 2016, AlphaGo beat Lee Sedol, 4:1

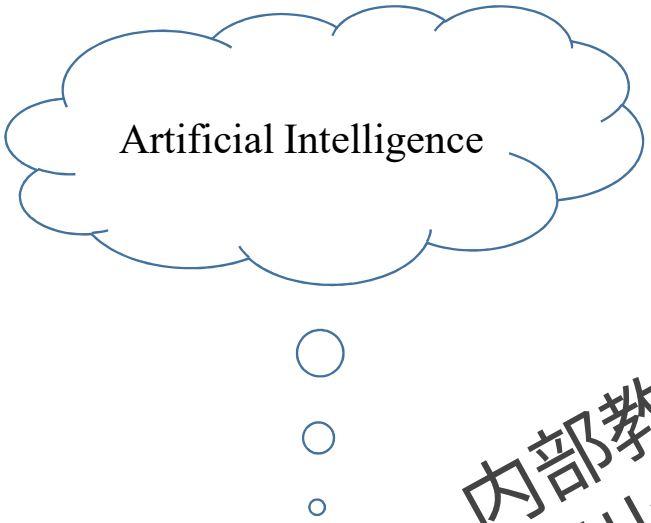
May 2017, AlphaGo beat Ke Jie, 3:0

Artificial Intelligence

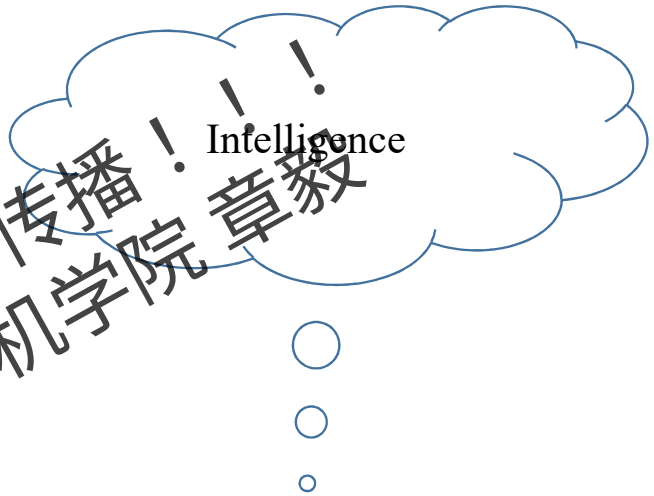
The main working principle is deep learning.



Concepts



Artificial ⊕ Intelligence

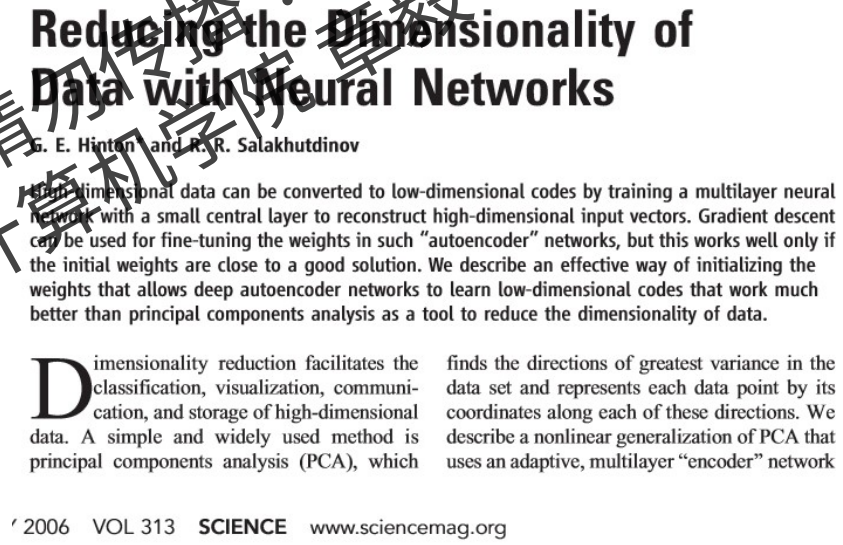


- *Definition of Intelligence*
- Intelligence is the ability to learn about, learn from, understand, and interact with one's environment.

Concepts

- *Deep Learning*
  - Method to train Deep Neural Networks

Professor Geoffrey E. Hinton



Reducing the Dimensionality of Data with Neural Networks

G. E. Hinton* and R. R. Salakhutdinov

High-dimensional data can be converted to low-dimensional codes by training a multilayer neural network with a small central layer to reconstruct high-dimensional input vectors. Gradient descent can be used for fine-tuning the weights in such "autoencoder" networks, but this works well only if the initial weights are close to a good solution. We describe an effective way of initializing the weights that allows deep autoencoder networks to learn low-dimensional codes that work much better than principal components analysis as a tool to reduce the dimensionality of data.

Dimensionality reduction facilitates the classification, visualization, communication, and storage of high-dimensional data. A simple and widely used method is principal components analysis (PCA), which finds the directions of greatest variance in the data set and represents each data point by its coordinates along each of these directions. We describe a nonlinear generalization of PCA that uses an adaptive, multilayer "encoder" network

2006 VOL 313 SCIENCE www.sciencemag.org

Concepts

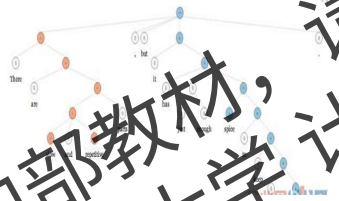
Applications

- Speech recognition：RBM network reduce the error rate by 30%, which is the most significant breakthrough in the past decade.
- Image recognition：CNN achieve over 95% recognition rate on Imagenet2012 dataset, which is comparable to human performance.
- Nature language processing：recurrent neural networks show superior performance than baseline methods in NLP problems..

Microsoft simultaneous interpretation

Products

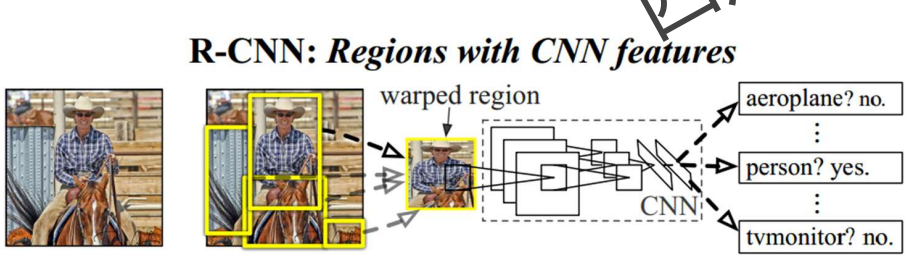
IMAGENET



Stanford sentiment analysis

Google translation

Google voice assistance

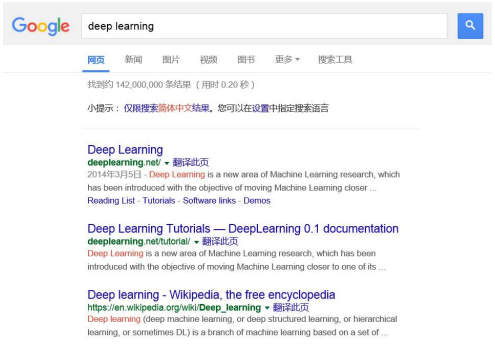


A group of young men playing a game of soccer.

A black and white cat is sitting on a chair.

Concepts

- An advanced technology in computer science



Deep learning heroes

Geoffrey Hinton
University of Toronto
Google

Yann LeCun
New York University
Facebook

Andrew Ng
Stanford University
Baidu

Yoshua Bengio
University of Montreal

Prof. Jürgen Schmidhuber
IDSIA Dalle Molle Institute for Artificial Intelligence

Scuola universitaria professionale della Svizzera italiana
SUPSI

Università della Svizzera italiana

Goodfellow

Baidu 百度

Google IBM Microsoft facebook

Deep learning is a new area of machine learning research, which has been introduced with the objective of moving machine learning closer to one of its original goals: artificial intelligence.

Concepts

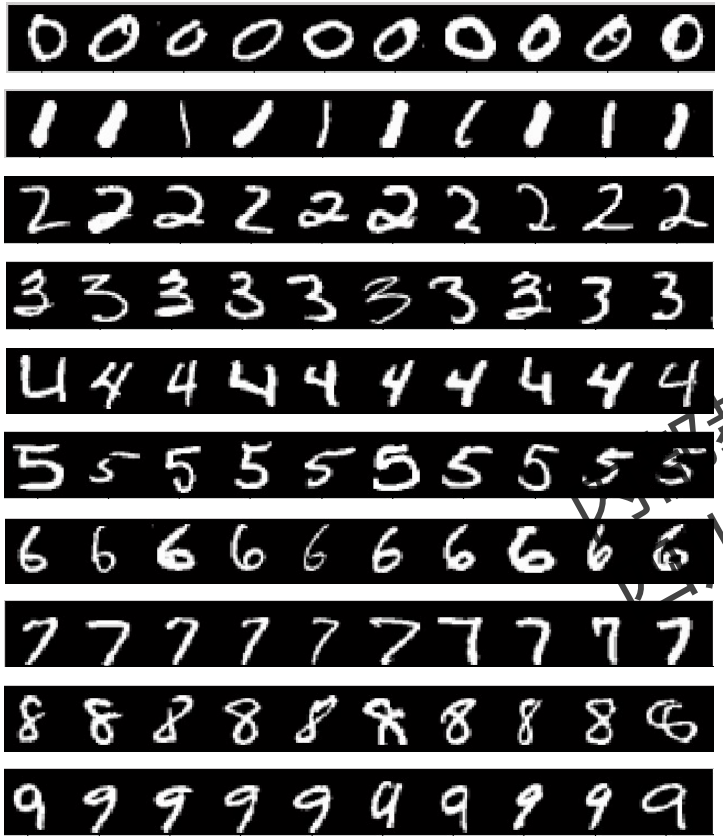
Deep Neural Networks
A way towards artificial intelligence

Deep Neural Networks are computational models inspired from brain working principle.

Outline

- Concepts
- An example: handwritten digits recognition
- How does a child recognize the handwritten digits
- Introduction to brain structure
- Discussions

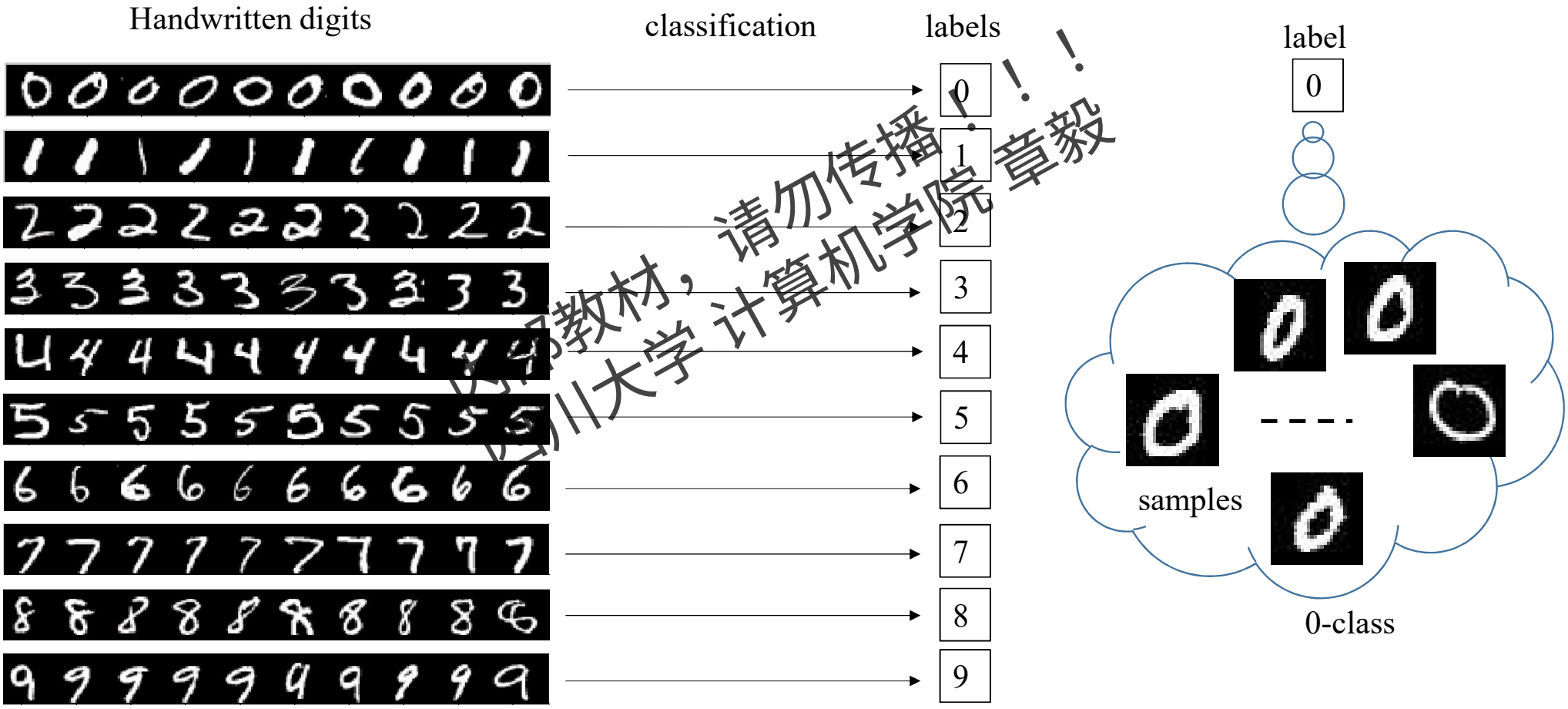
Handwritten Digits Recognition



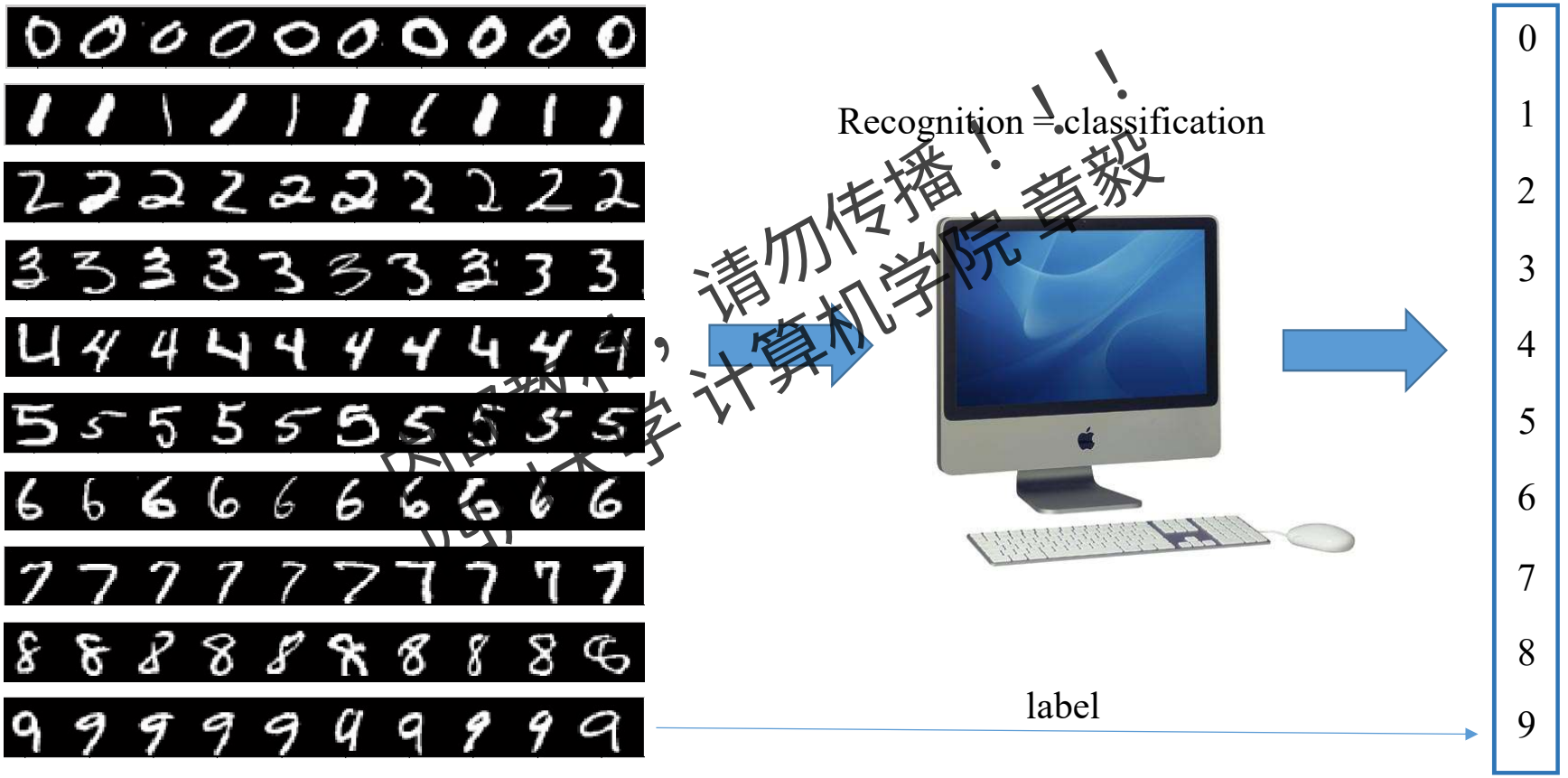
Problem:
How to write a computer program to recognize the handwritten digits?

?

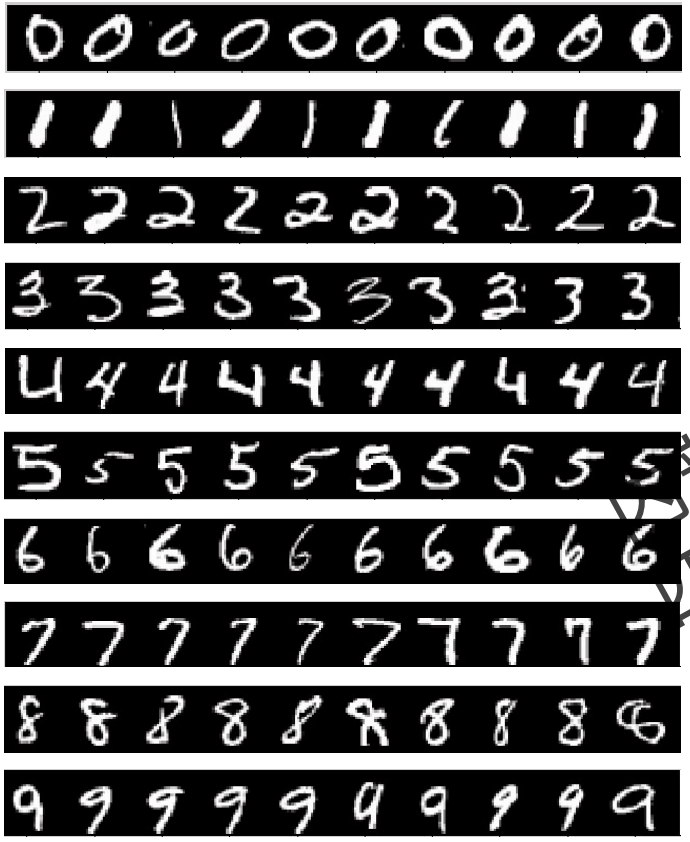
Handwritten Digits Recognition



Handwritten Digits Recognition



Handwritten Digits Recognition



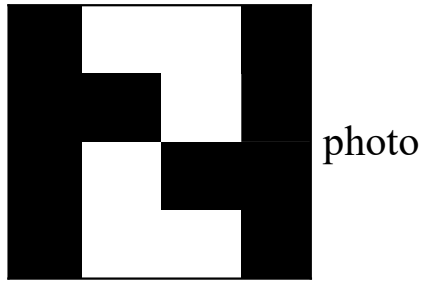
I cannot read picture directly.

digitizing

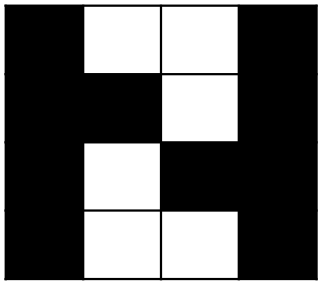
representation

0
1
2
3
4
5
6
7
8
9

Handwritten Digits Recognition

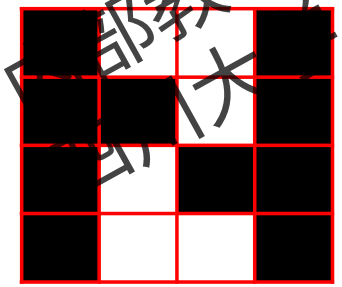


photo



4×4

digitizing



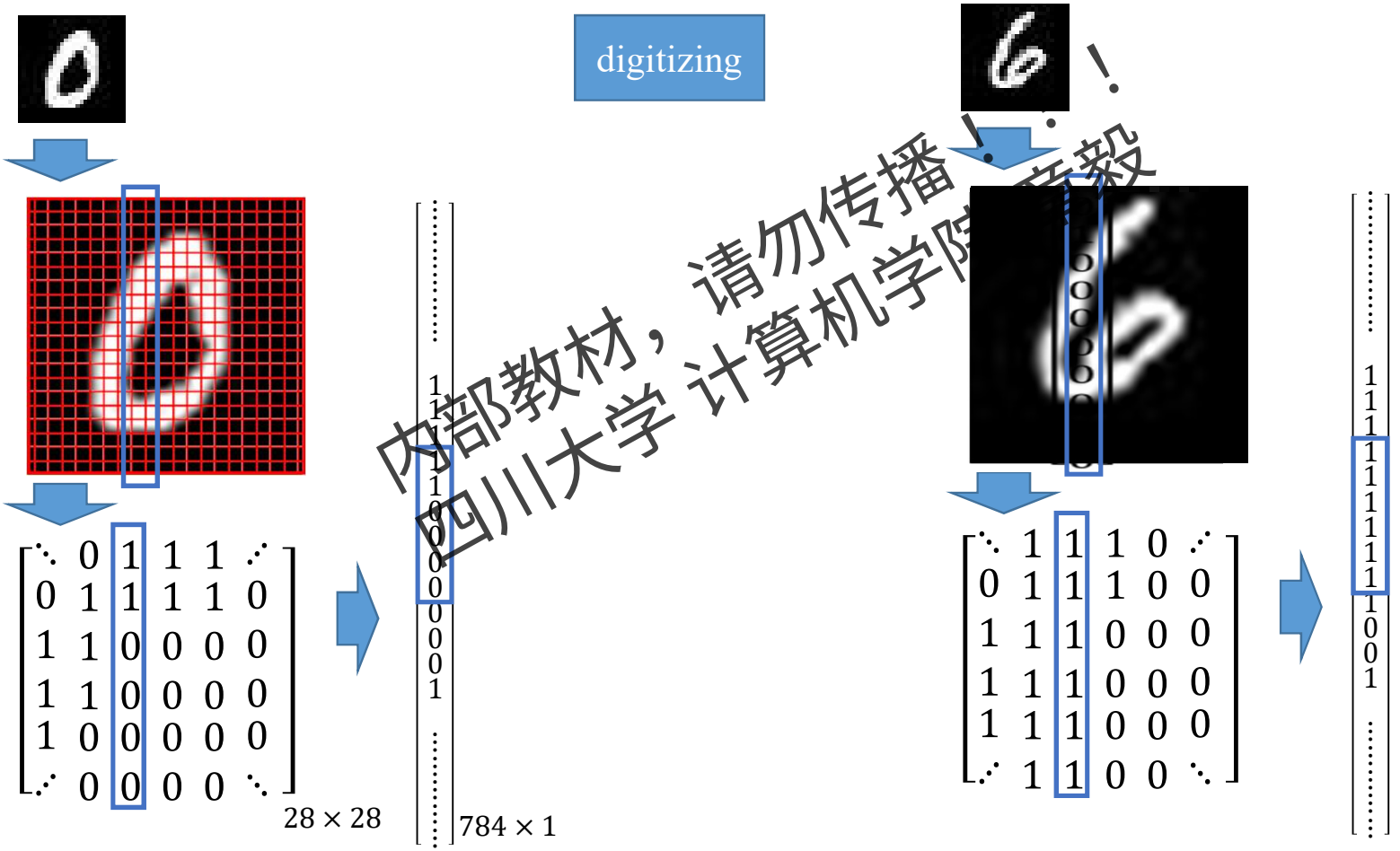
$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

4×4

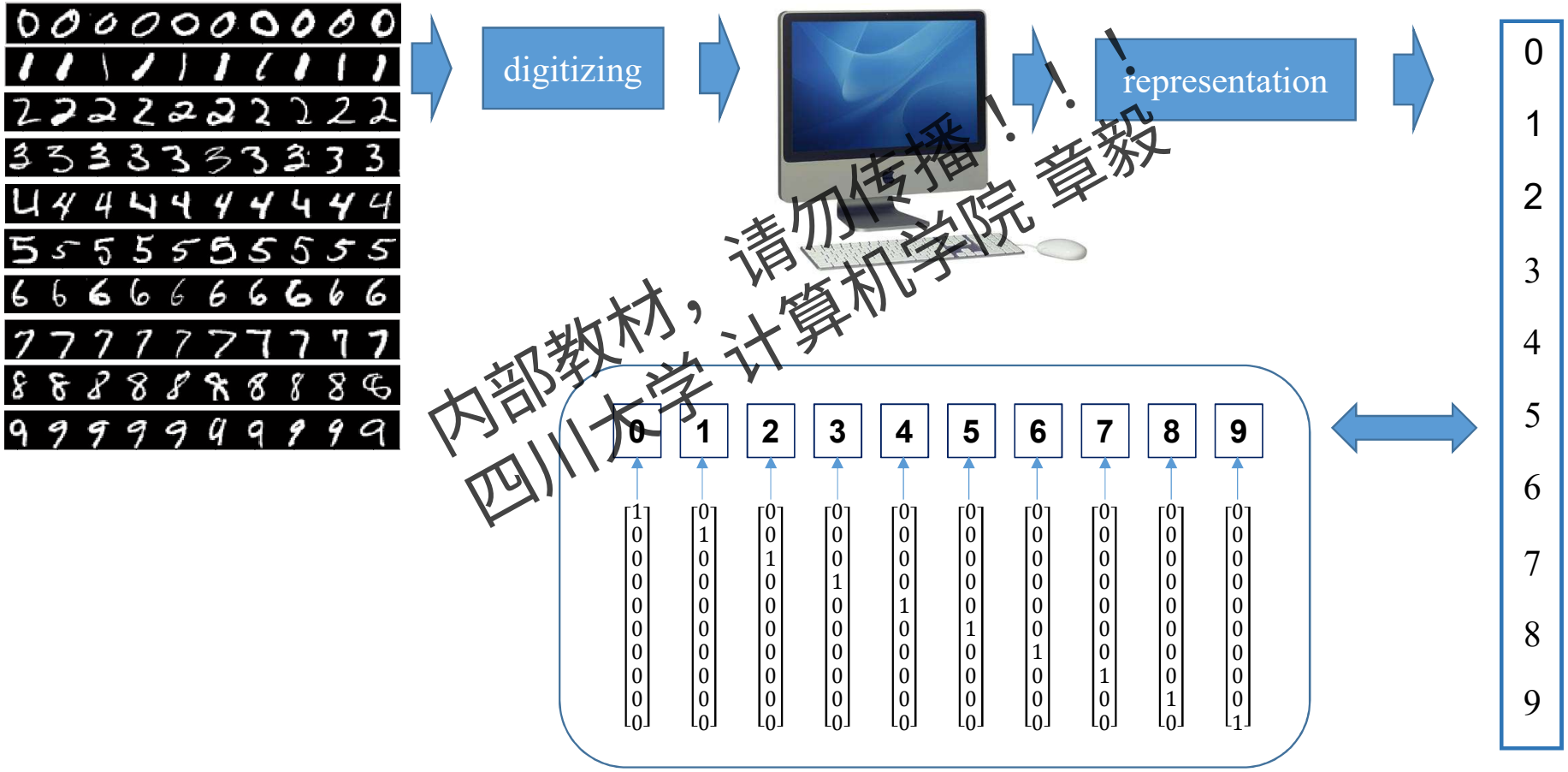
$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

16×1

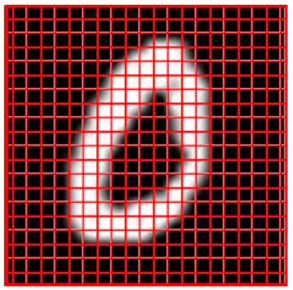
Handwritten Digits Recognition



Handwritten Digits Recognition



Handwritten Digits Recognition



pixels

28×28

matrix

$$\begin{bmatrix} \ddots & 0 & 1 & 1 & 1 & \vdots \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ \cdots & 0 & 0 & 0 & 0 & \ddots \end{bmatrix}$$

28×28

vector

$$x = \begin{bmatrix} \vdots \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ \vdots \end{bmatrix}$$

784×1

$$y = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

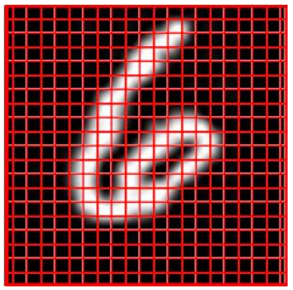
0

Handwritten Digits Recognition

$$y = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

6

pixels



28×28

matrix

$$\begin{bmatrix} \ddots & 1 & 1 & 1 & 0 & \vdots \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ \cdots & 1 & 1 & 0 & 0 & \ddots \end{bmatrix}$$

28×28

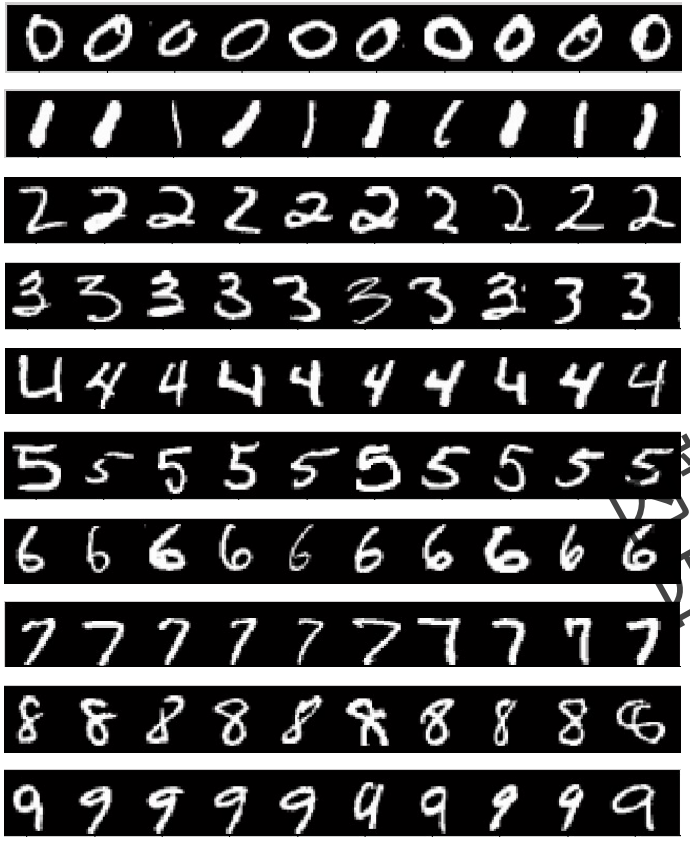
vector

$$x = \begin{bmatrix} \vdots \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ \vdots \end{bmatrix}$$

784×1

Problem:
How can I correctly output the representation associated with the input?

Handwritten Digits Recognition



$F: x \rightarrow y$

x

Input sample

y

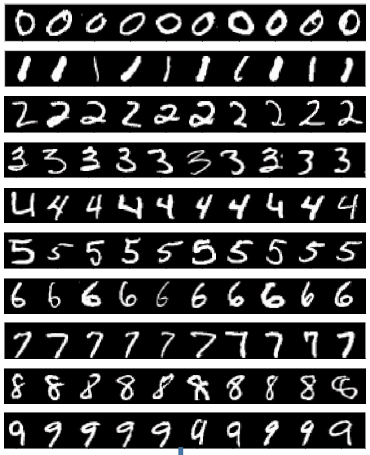
Label representation

$y = F(x)$

Are there such F exist?

0
1
2
3
4
5
6
7
8
9

Handwritten Digits Recognition



$x =$

$$\begin{bmatrix} \vdots \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ \vdots \end{bmatrix}$$

28x28=784

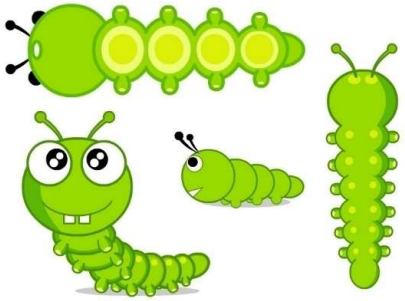
$F: x \rightarrow y$

F is map, it maps from 784 dimensional space to 10 dimensional space.

$y =$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

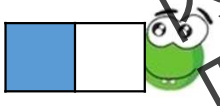
An example: XOR-worms problem



内部教材，请勿传播！！！
四川大学 计算机学院 章毅

Can you read us?

Doted worms

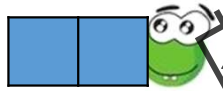


Smooth worms

An example: XOR-worms problem

Doted worms

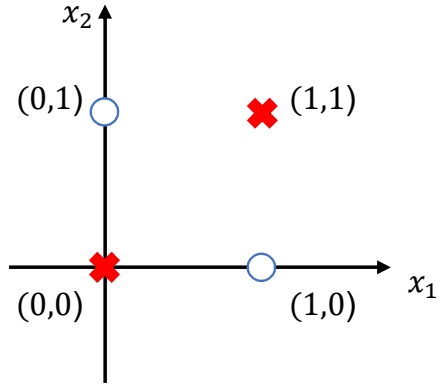
Smooth worms



1

0

An example: XOR-worms problem



Doted worms

$\begin{bmatrix}0\\1\end{bmatrix}$

$\begin{bmatrix}1\\0\end{bmatrix}$

1

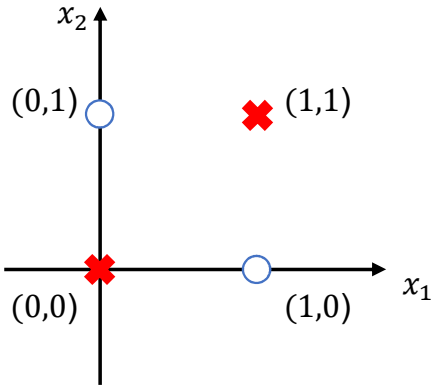
Smooth worms

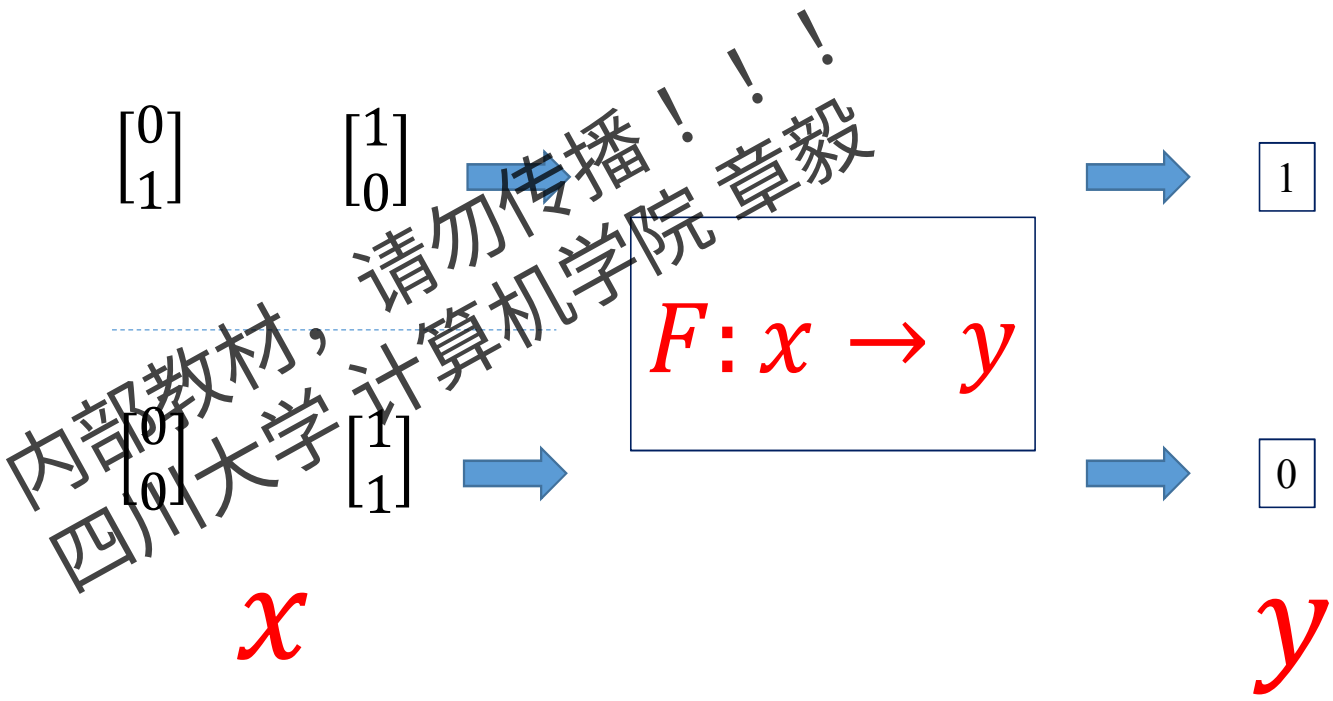
$\begin{bmatrix}0\\0\end{bmatrix}$

$\begin{bmatrix}1\\1\end{bmatrix}$

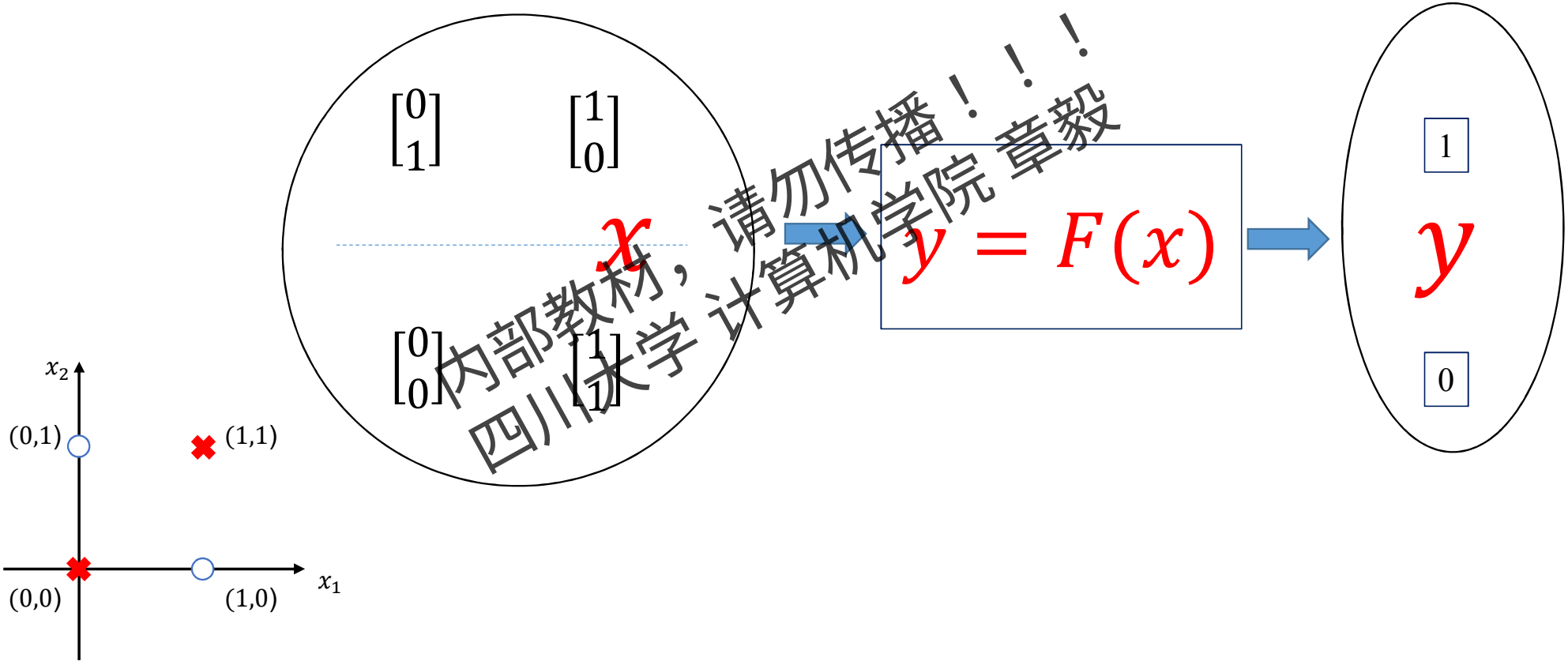
0

An example: XOR-worms problem

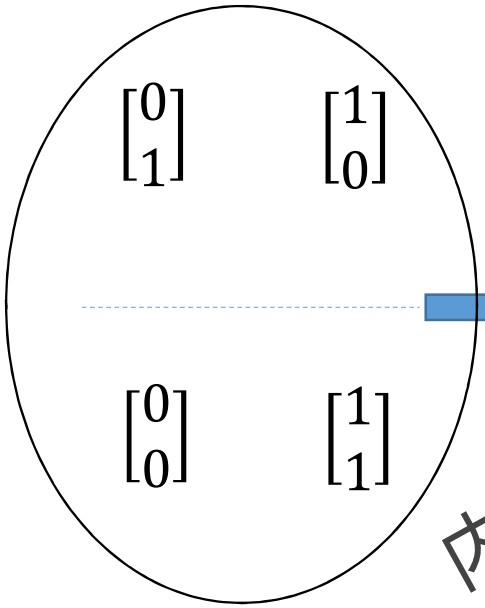




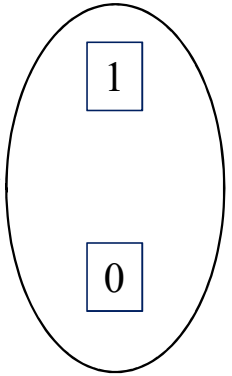
An example: XOR-worms problem



An example: XOR-worms problem

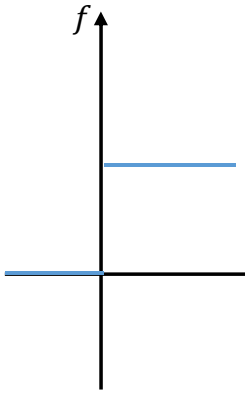


$$F\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = f[f(2x_1 + 2x_2 - 1) + f(-x_1 - x_2 + 1.5) - 1.5]$$



$$F\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = f[f(2\cdot 0 + 2\cdot 1 - 1) + f(-0 - 1 + 1.5) - 1.5] = 1$$

$$F\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = f[f(2\cdot 1 + 2\cdot 1 - 1) + f(-1 - 1 + 1.5) - 1.5] = 0$$



$$f(s) = \begin{cases} 1, & s \geq 0 \\ 0, & otherwise \end{cases}$$

Problem:
How to construct the mapping F?

An example: XOR-worms problem

Doted worms

$\begin{bmatrix}0\\1\end{bmatrix}$

$\begin{bmatrix}1\\0\end{bmatrix}$

Smooth worms

$\begin{bmatrix}0\\0\end{bmatrix}$

$\begin{bmatrix}1\\1\end{bmatrix}$

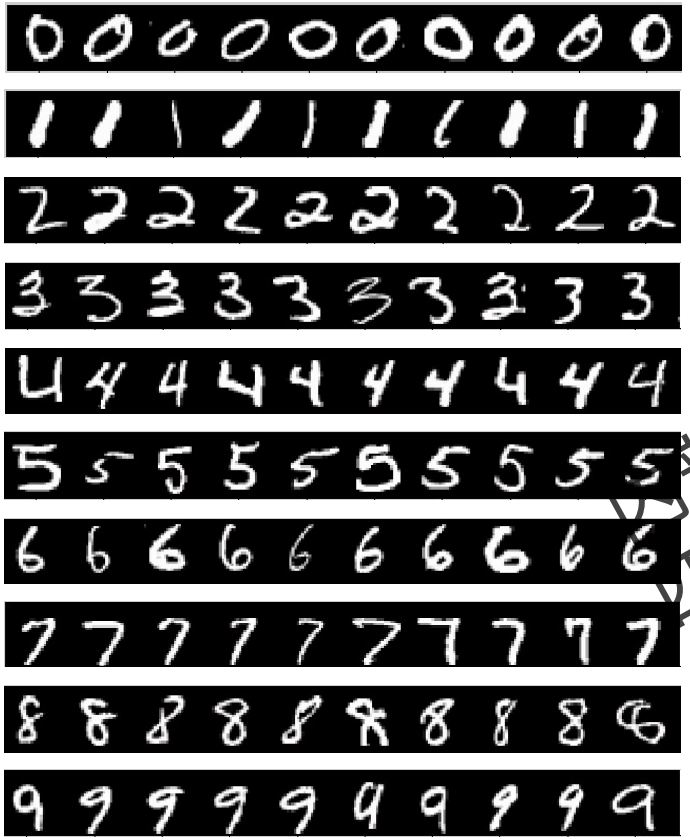
$$F\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = f[f(2x_1 + 2x_2 - 1) + f(-x_1 - x_2 + 1.5) - 1.5]$$

1

0

Problem:
How to construct the mapping F?

Handwritten Digits Recognition



digitizing

x

$F: x \rightarrow y$

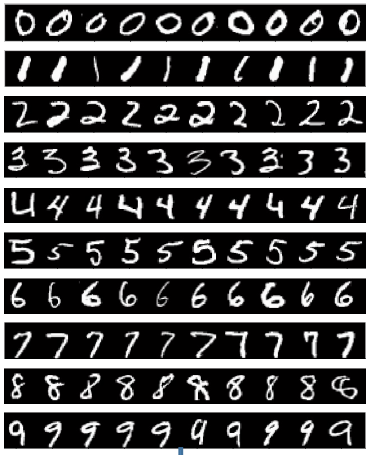
representation

y

0
1
2
3
4
5
6
7
8
9

Problem:
How to construct the mapping F?

Handwritten Digits Recognition



$x =$

$$\begin{bmatrix} \vdots \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ \vdots \end{bmatrix}$$

28x28=784

$$y = F(x)$$

$y =$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

内部教材，请勿传播！！！
四川大学 计算机学院 黄毅

Problem:

How to construct the mapping F from 784 dimensional space to 10 dimensional space?

It is almost impossible manually!

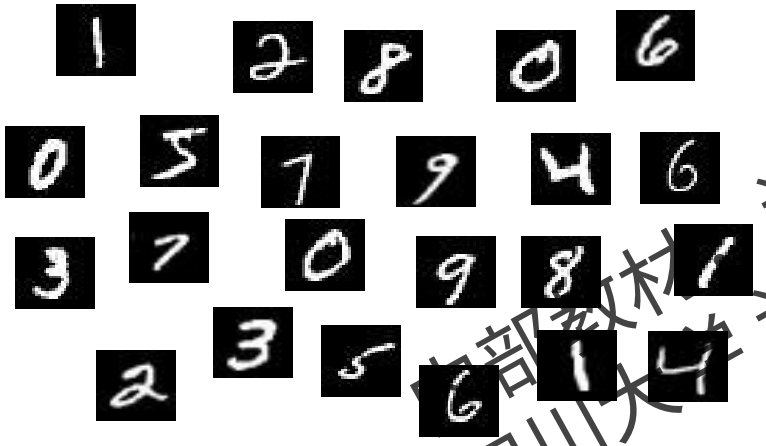
Any other methods? Next, let's see how a child can easily do it.

Outline

- Concepts
- An example: handwritten digits recognition
- How does a child recognize the handwritten digits
- Introduction to brain structure
- Discussions

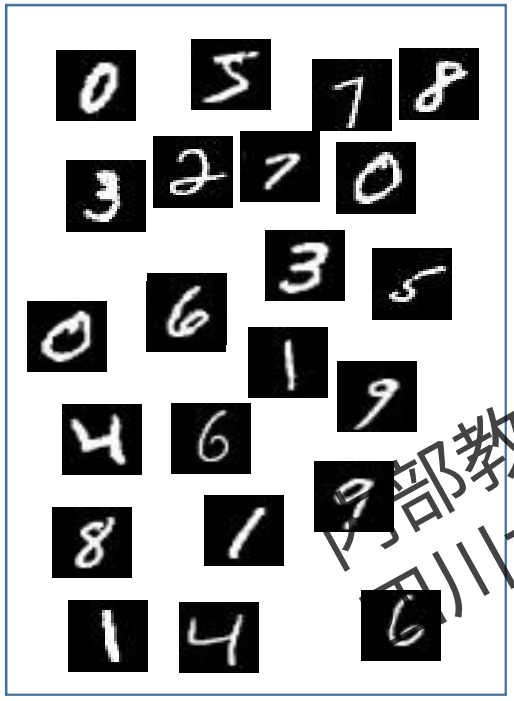
How does a child recognize the handwritten digits?

How does a child recognize the handwritten digits?



Training

How does a child recognize the handwritten digits?



Training

Training Data

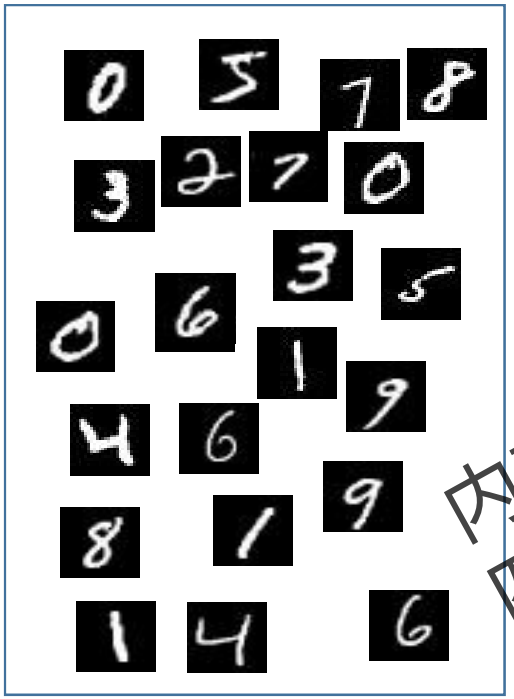
Training Data is a set of samples used during training.
Each sample is a pair of digit image and its label.

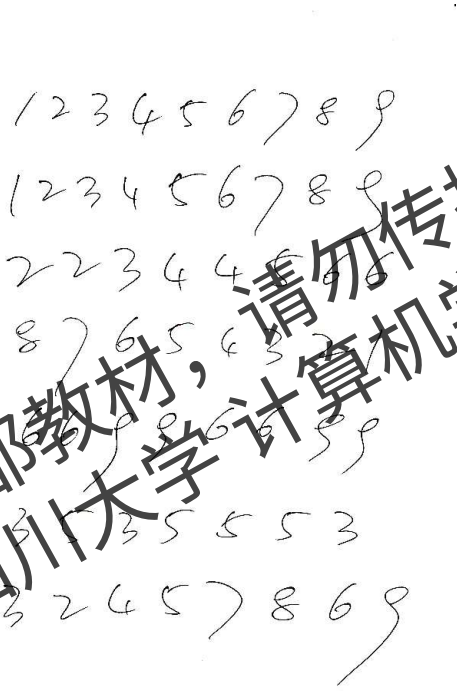
How does a child recognize the handwritten digits?

Using Training Data for testing

Under-fitting: perform badly on training data

How does a child recognize the handwritten digits?



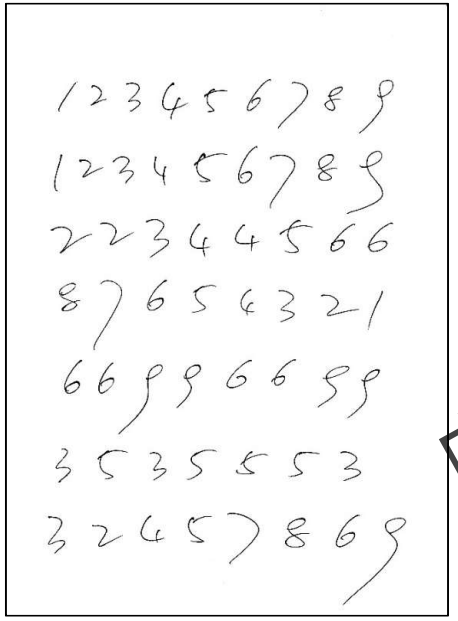


Training

Training Data

Training Data is a set of samples used during training.
Each sample is a pair of digit image and its label.

How does a child recognize the handwritten digits?



Test Data

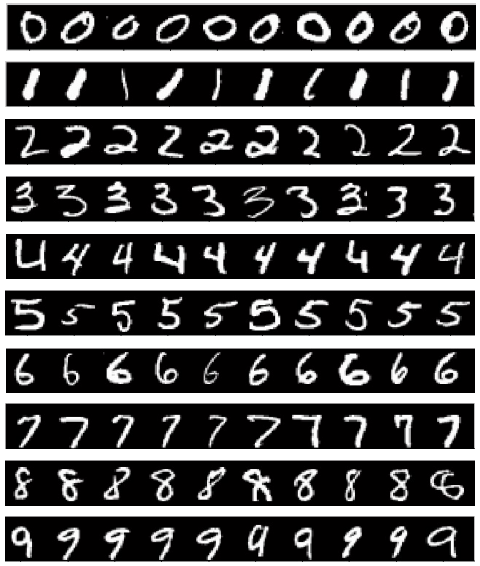
Test Data is a set of samples used for testing.
Each sample is a pair of digit image and its label.
Note that, it is different from the training data.

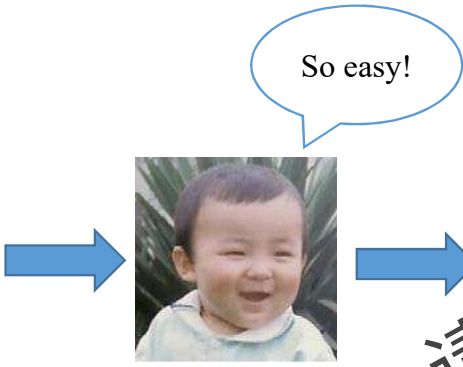
Testing

Over-fitting: perform very well on training data, but badly on test data

Generation performance: the performance on new data that hasn't been used during training

How does a child recognize the handwritten digits?





0
1
2
3
4
5
6
7
8
9

The human brain is so powerful so that any child can recognize the handwritten digits easily.
Two important factors:

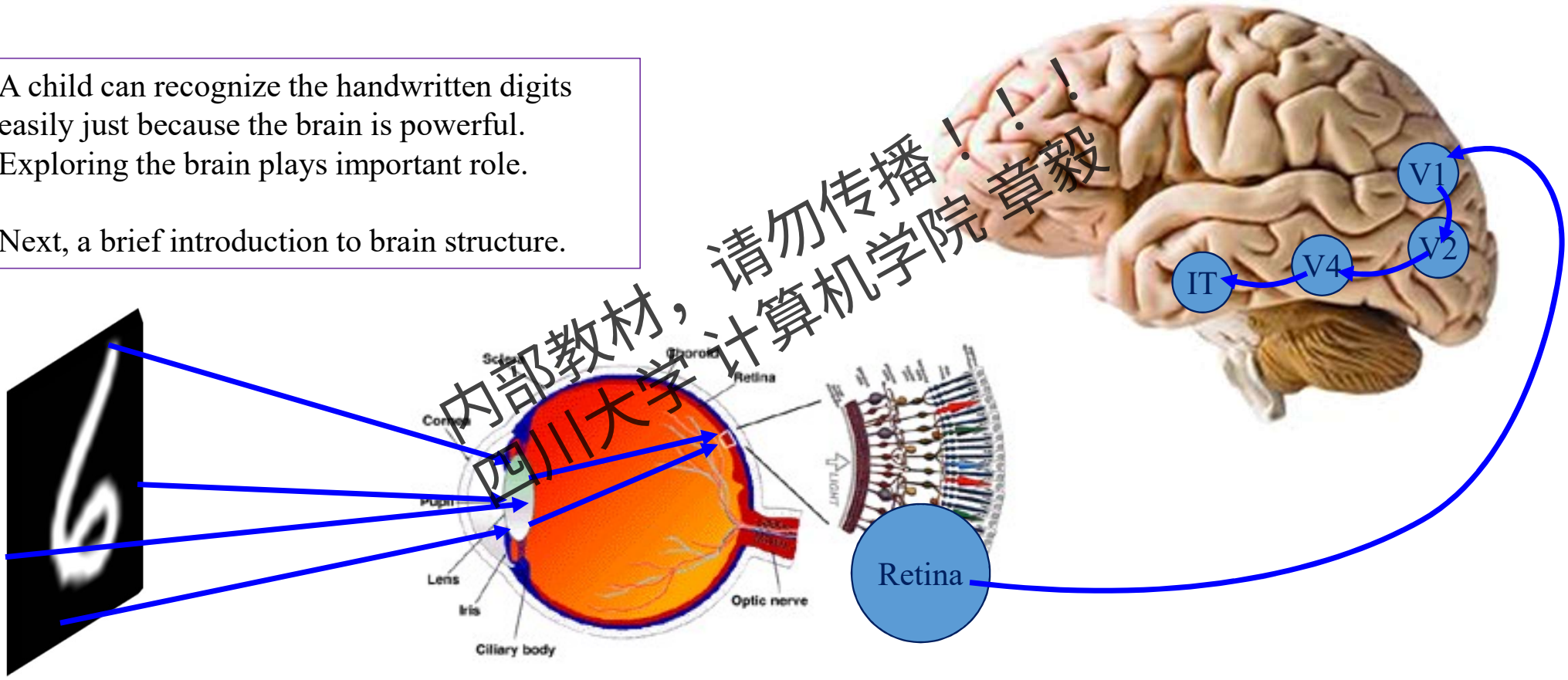
1. The brain has the structured ability.
2. Trained by some one.

Problem: How to develop methods for recognition by exploring the brain?

How does a child recognize the handwritten digits?

A child can recognize the handwritten digits easily just because the brain is powerful. Exploring the brain plays important role.

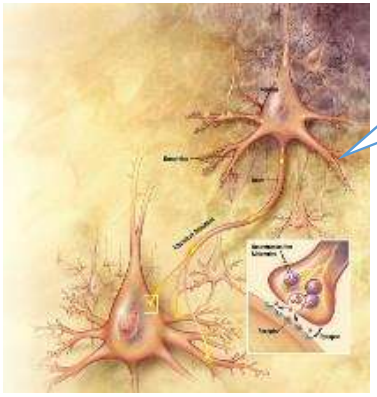
Next, a brief introduction to brain structure.



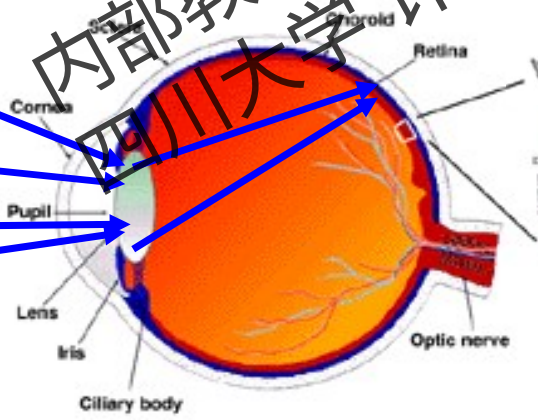
Outline

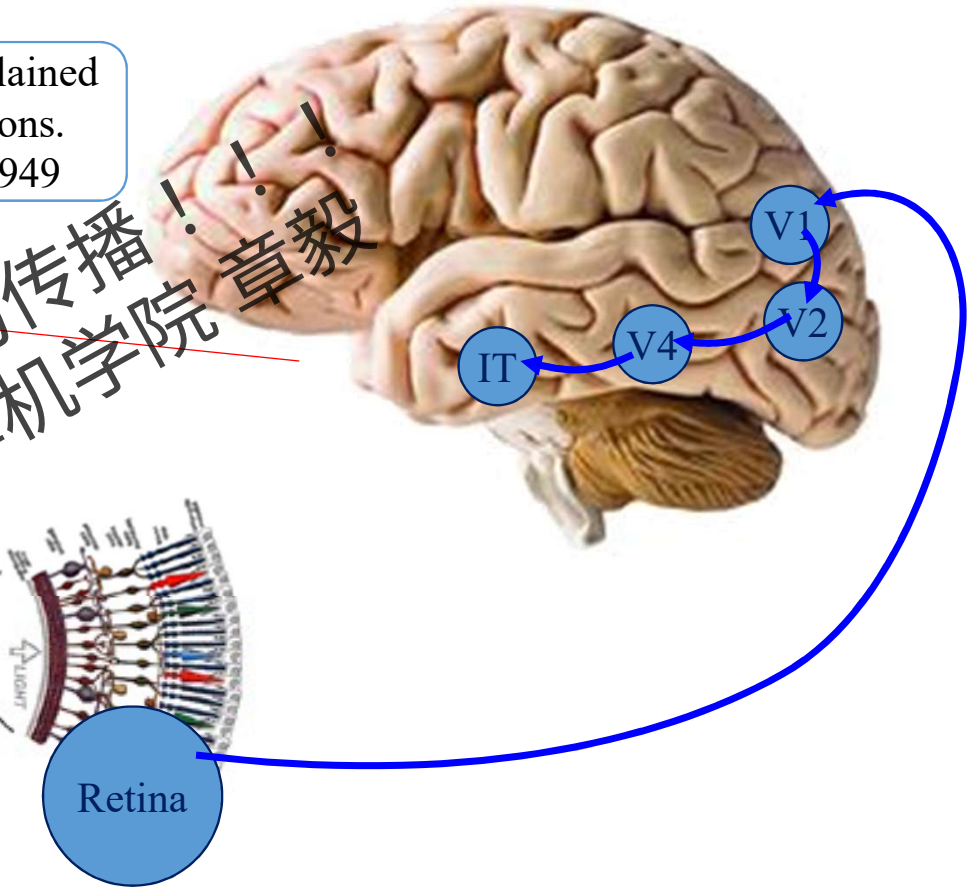
- Concepts
- An example: handwritten digits recognition
- How does a child recognize the handwritten digits
- Introduction to brain structure
- Discussions

Introduction to brain structure



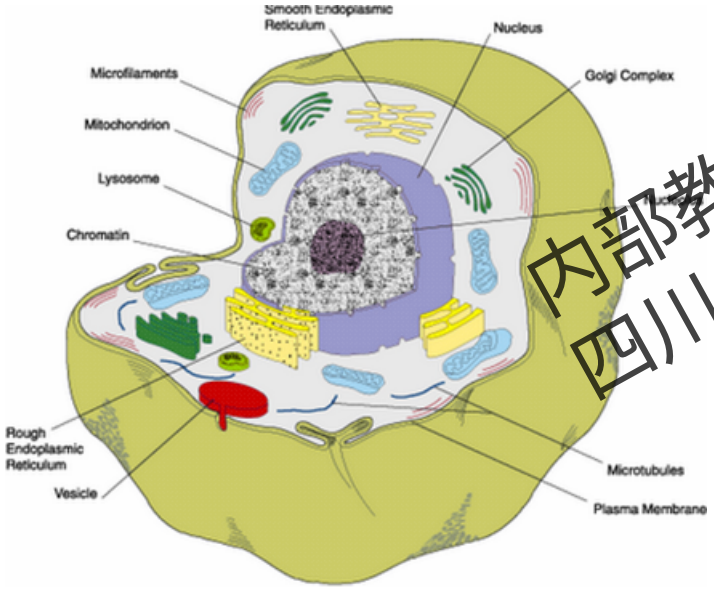
Behavior could be explained by the action of neurons.
------- D. O. Hebb, 1949



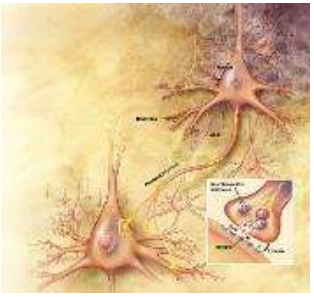


Introduction to brain structure

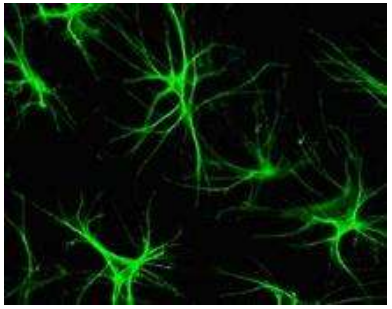
- All tissues and organs in the body consist of cells.



- Cells in the nervous system: neurons and glia.



neurons

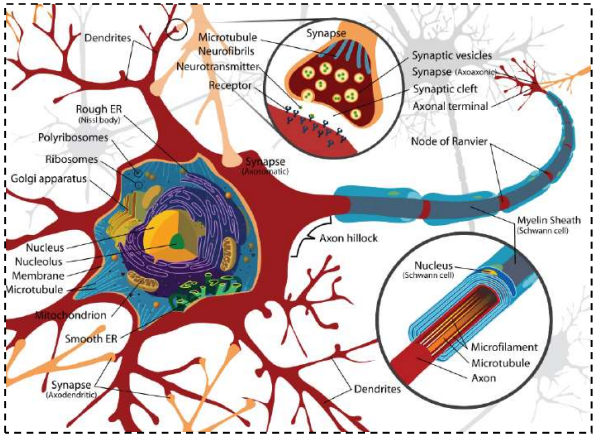


glia

Introduction to brain structure

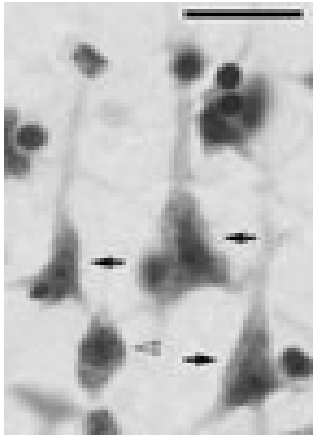
Retina resolution: 0.1mm

Neuron size: 0.01~0.05mm

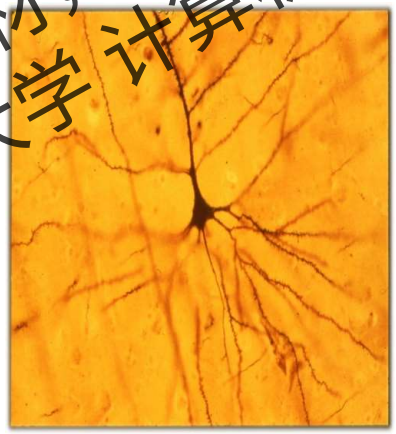


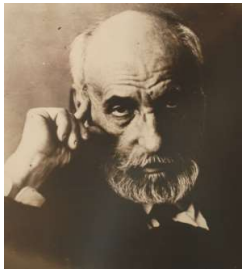
Introduction to brain structure

Franz Nissl
Germen, 19th century

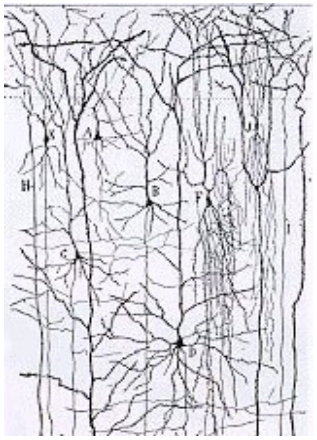


Camillo Golgi
Italian, 19th century

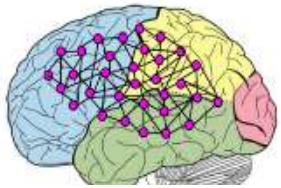


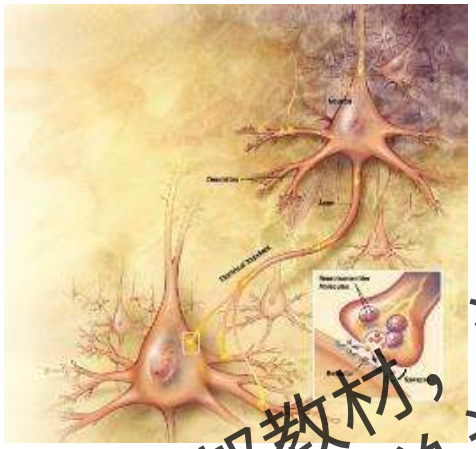


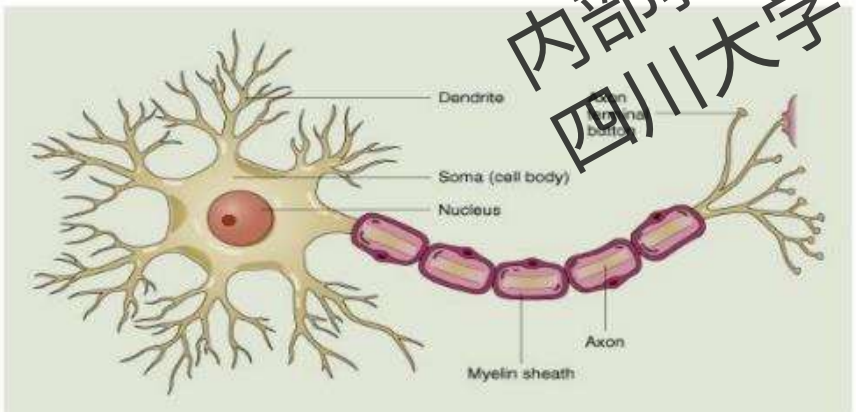
Santiago Ramón Cajal
Spanish, 19th century



Introduction to brain structure







What is a neuron?

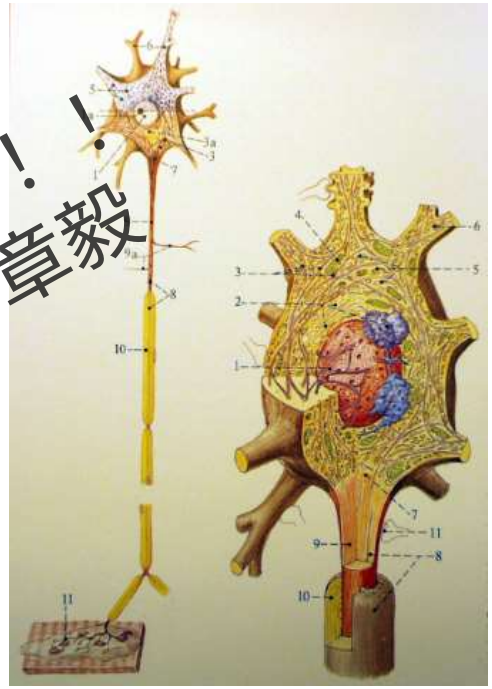
- Three main components
  - Soma (cell body)
  - Dendrite
  - Axon

- A neuron is an electrically excitable cell in brain.
- Neuron processes and transmits information through electrical and chemical signals.

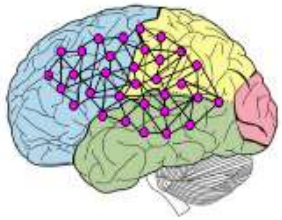
Introduction to brain structure

Brief statement of neuron

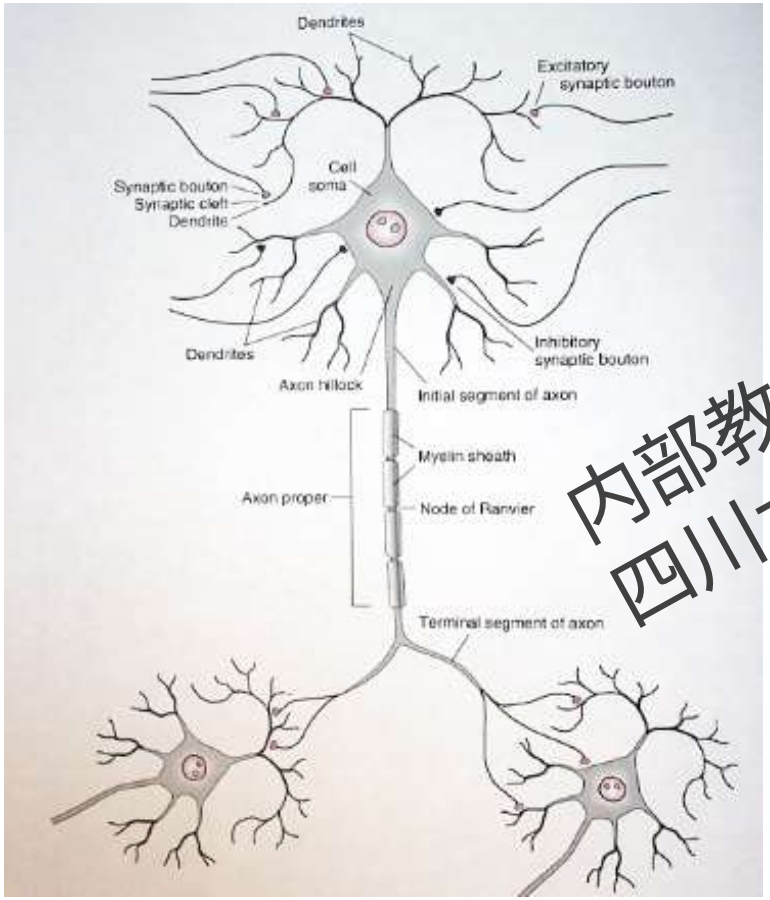
- Dendrites (inputs): receive chemical signals from other neurons.
- Soma (processing): collect and transfer and electrical signals.
- Axon (output): output signal to other neurons.



- A brain contains about 10^{11} neurons

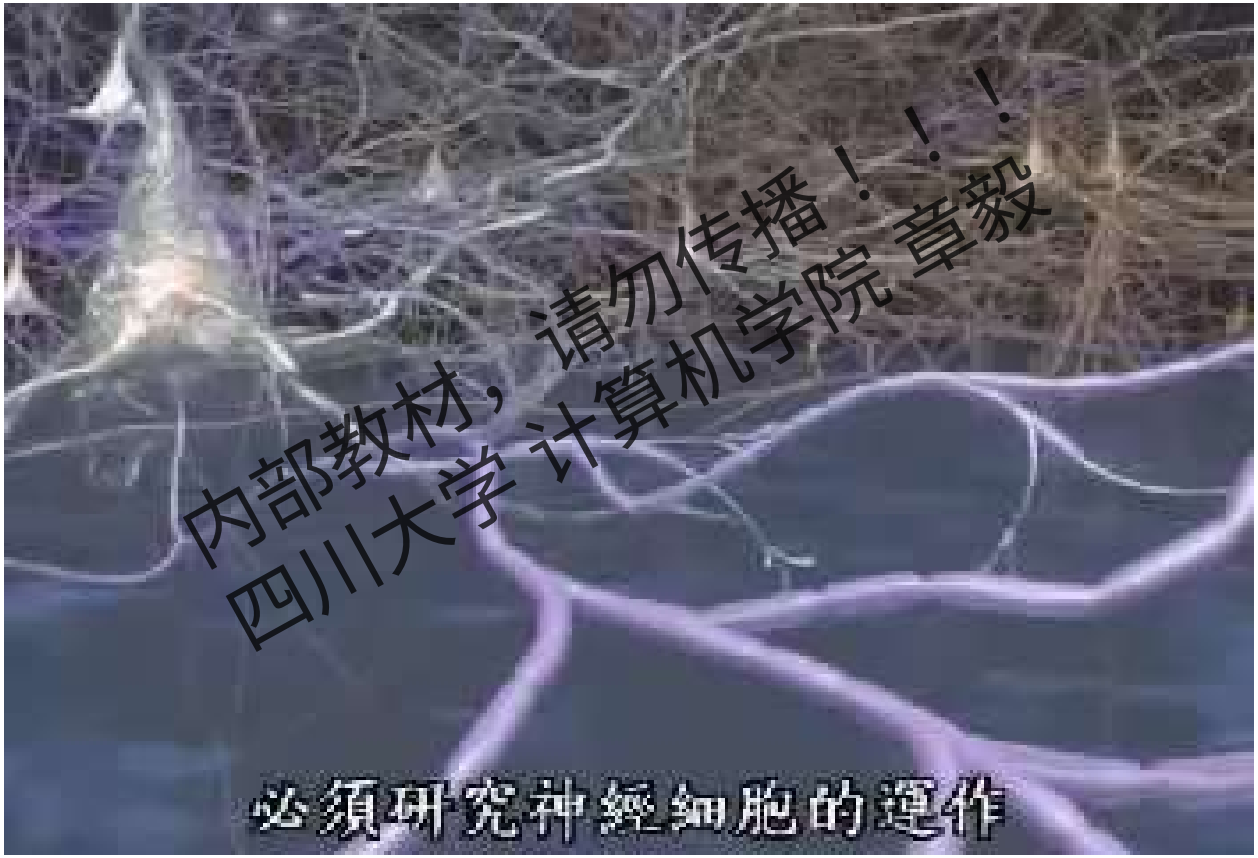


Introduction to brain structure



Neural Network = Neurons + Connections

Introduction to brain structure

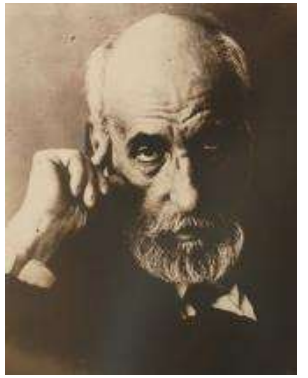


Introduction to brain structure

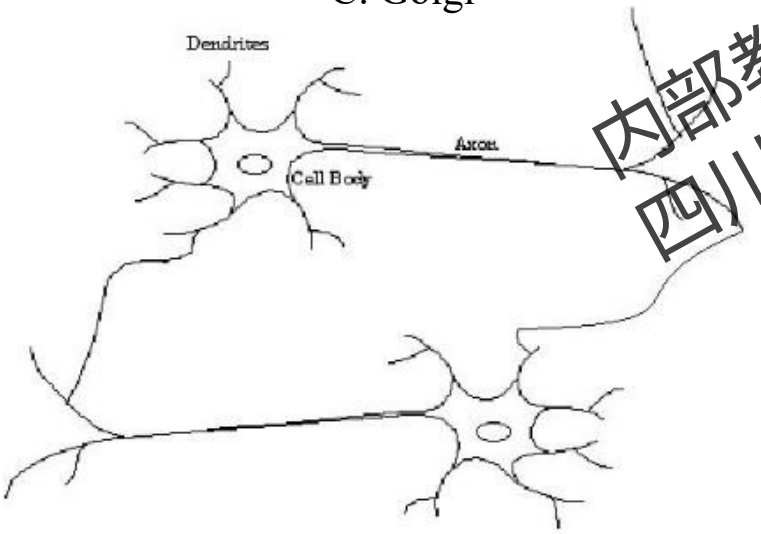
C. Golgi

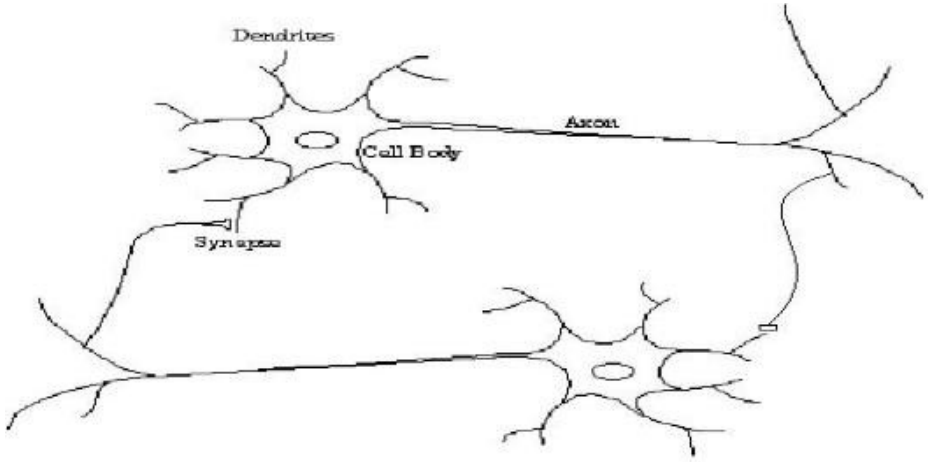
In what way are the neurons connected?

Shared the Nobel Prize in 1906

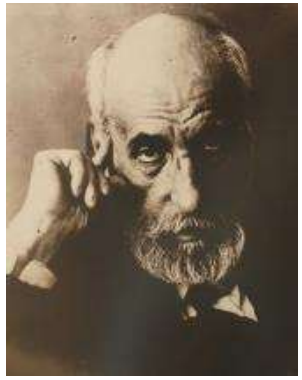


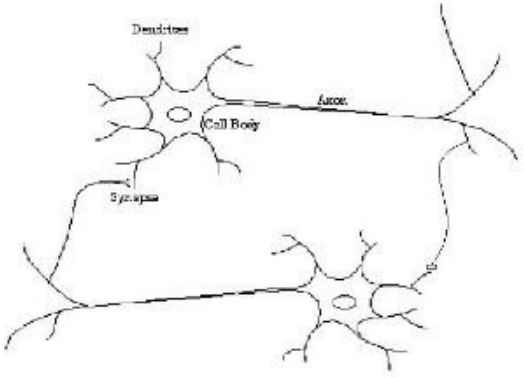
S. R. Cajal

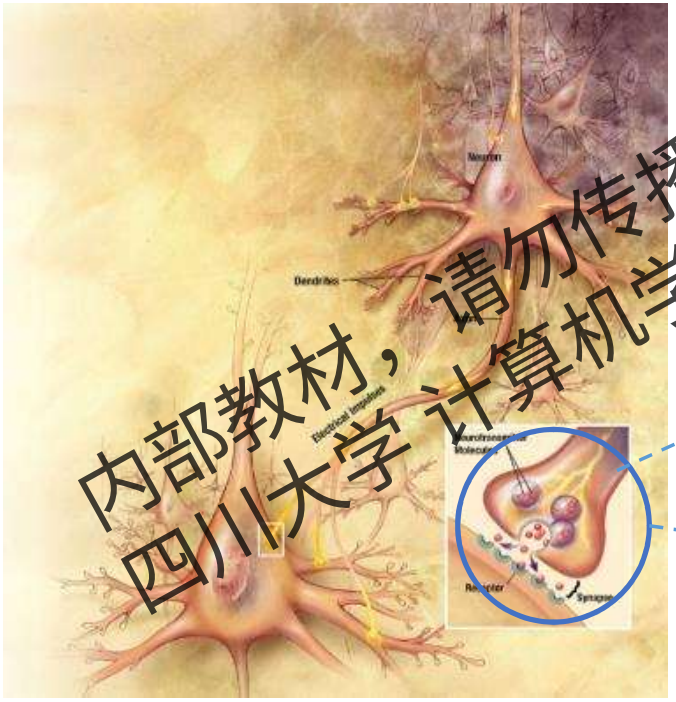




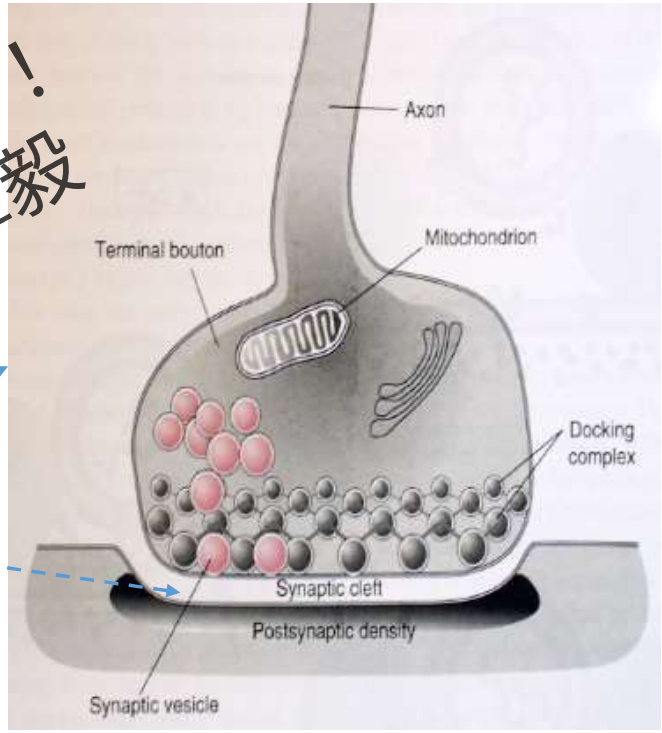
Introduction to brain structure





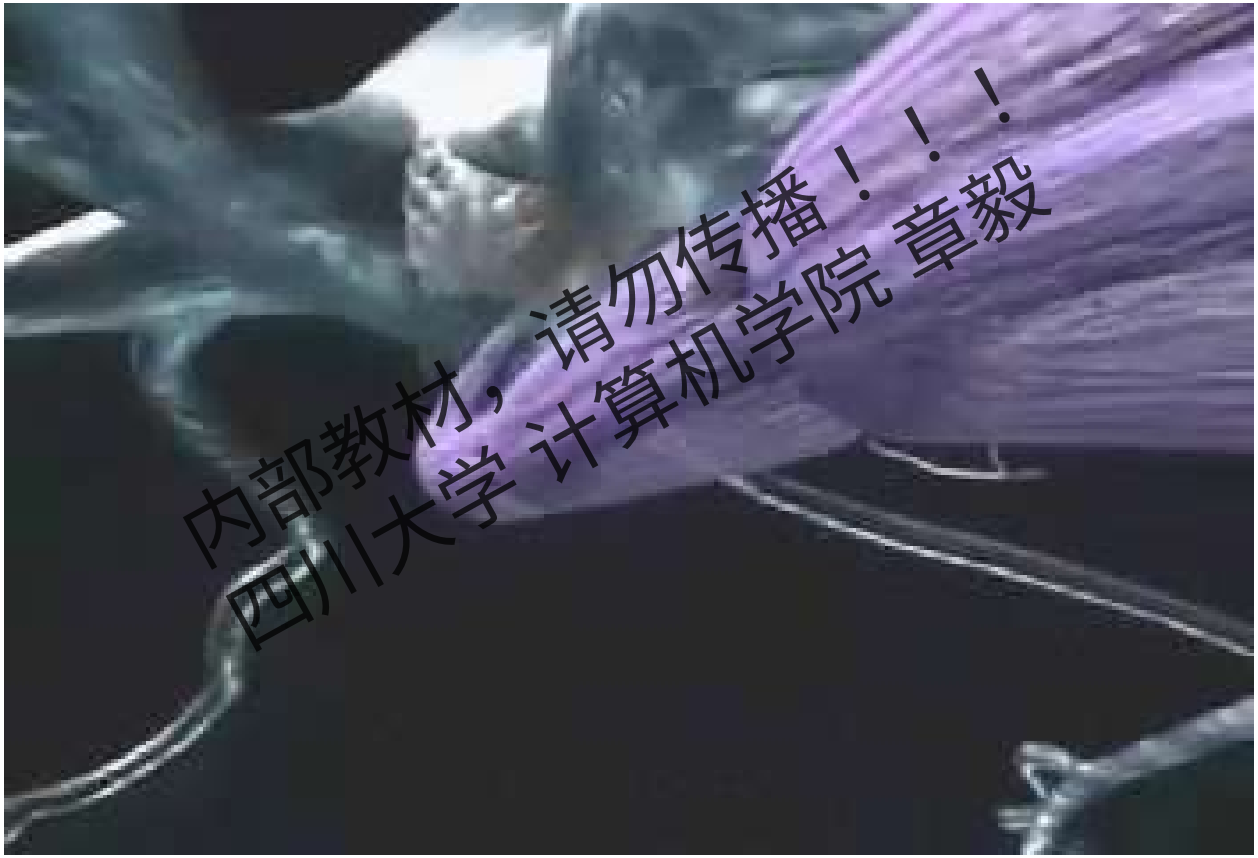


Synapse

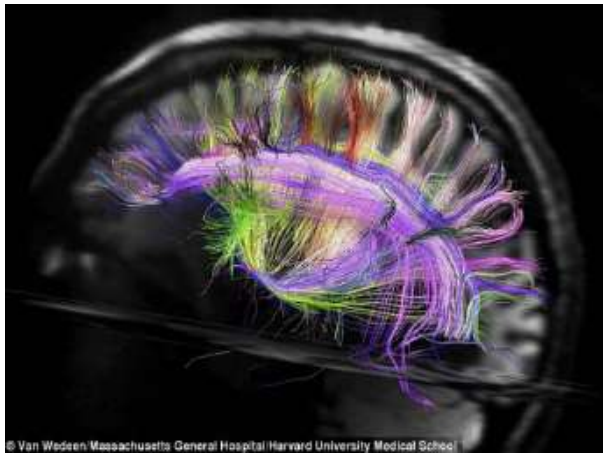


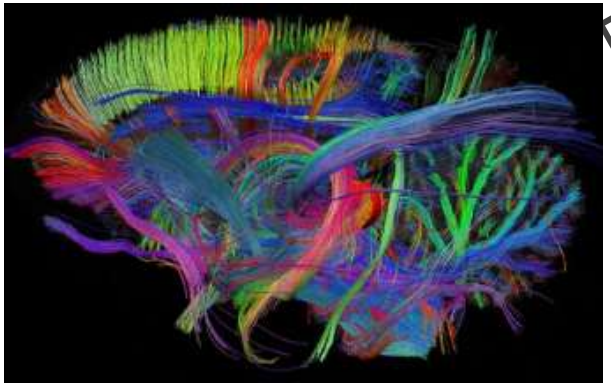
- A brain contains about 10^{11} neurons
- Each neuron has about 10^{4} connections

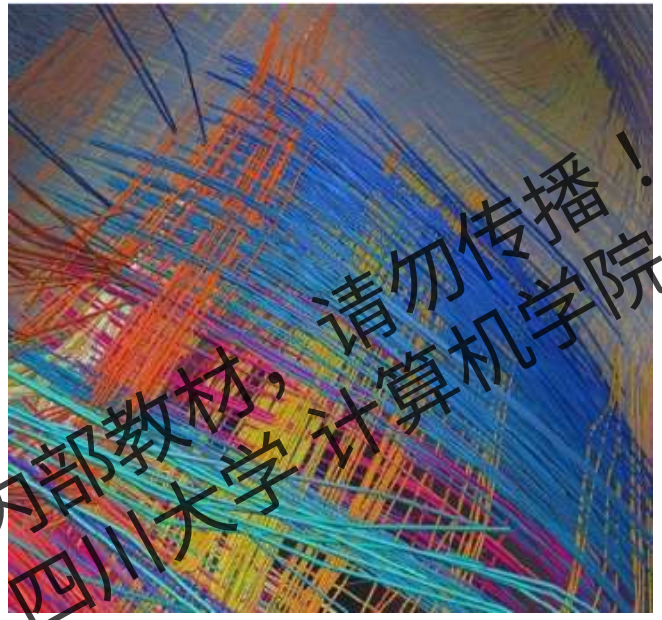
Introduction to brain structure

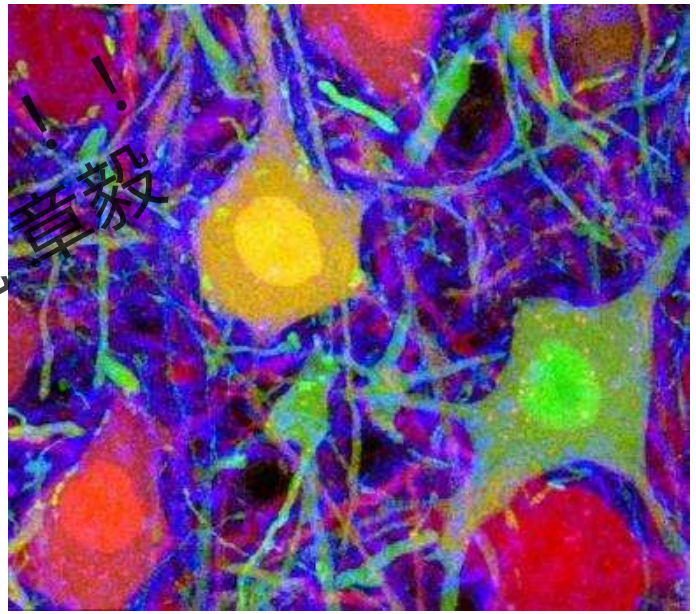


Introduction to brain structure

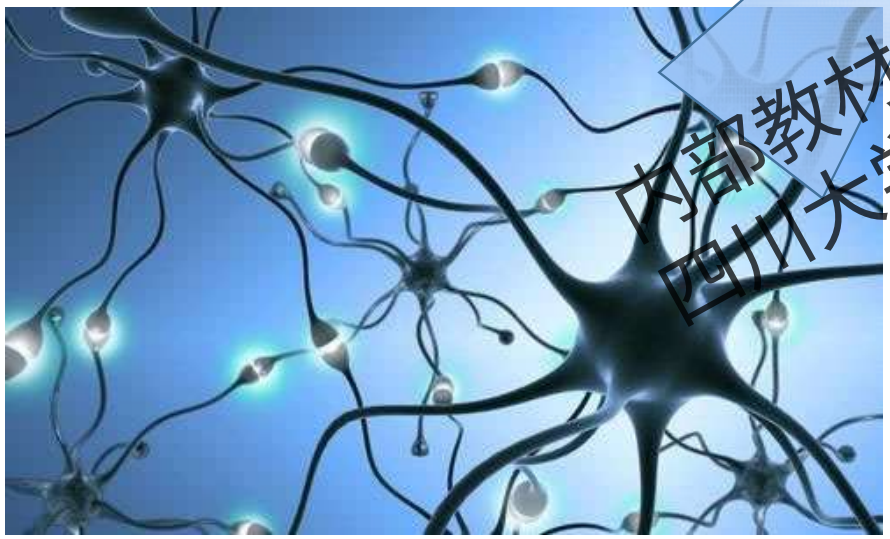


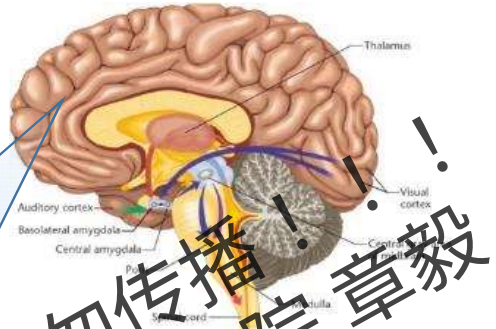


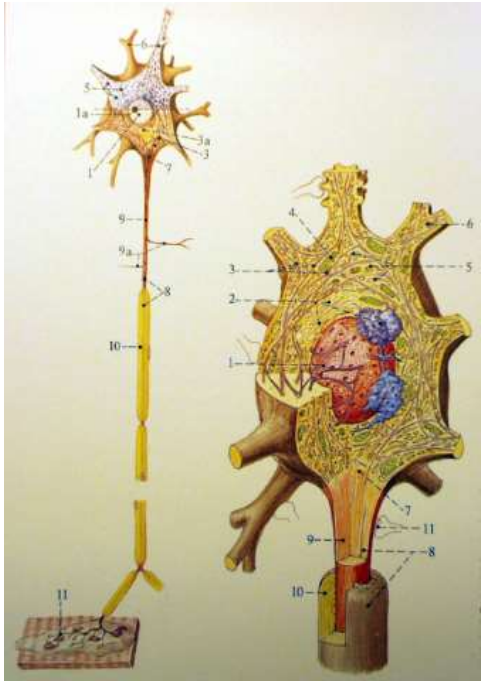




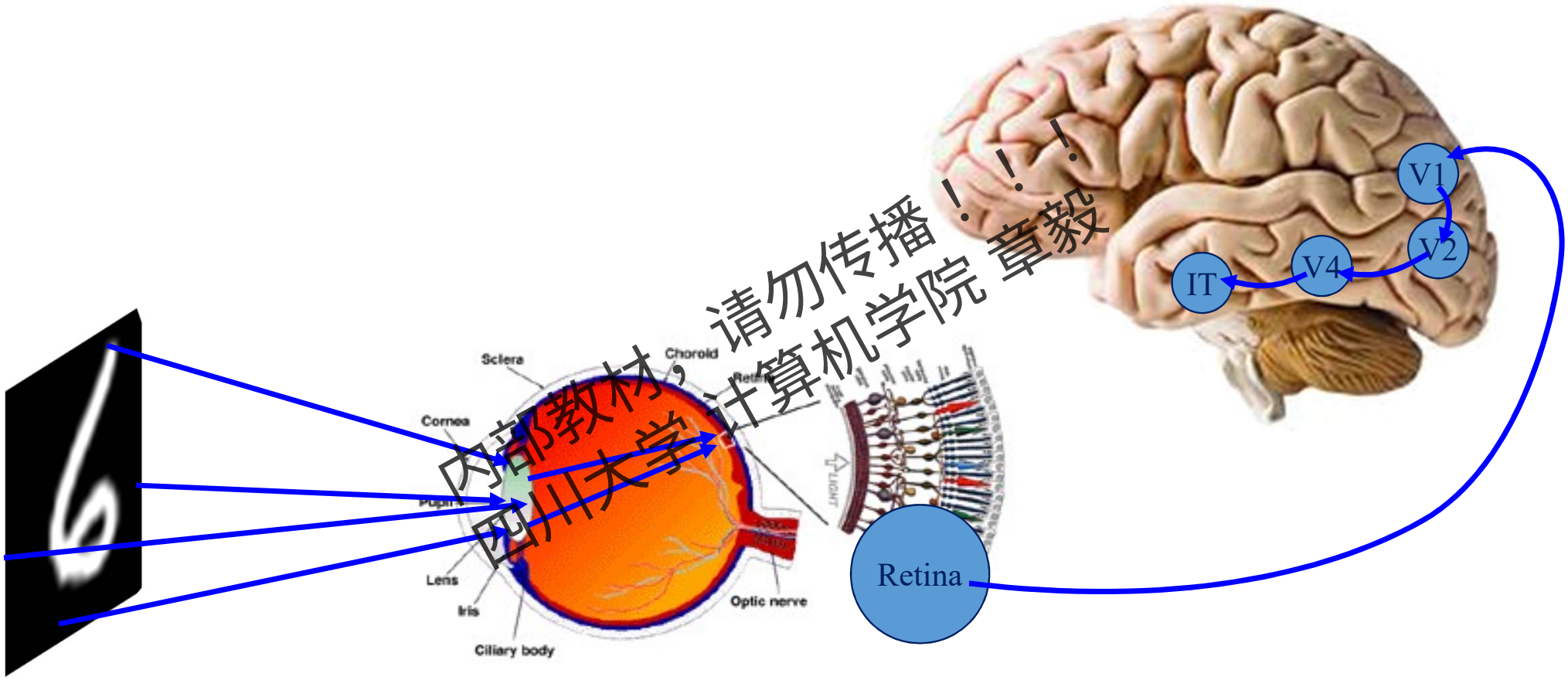
Introduction to brain structure



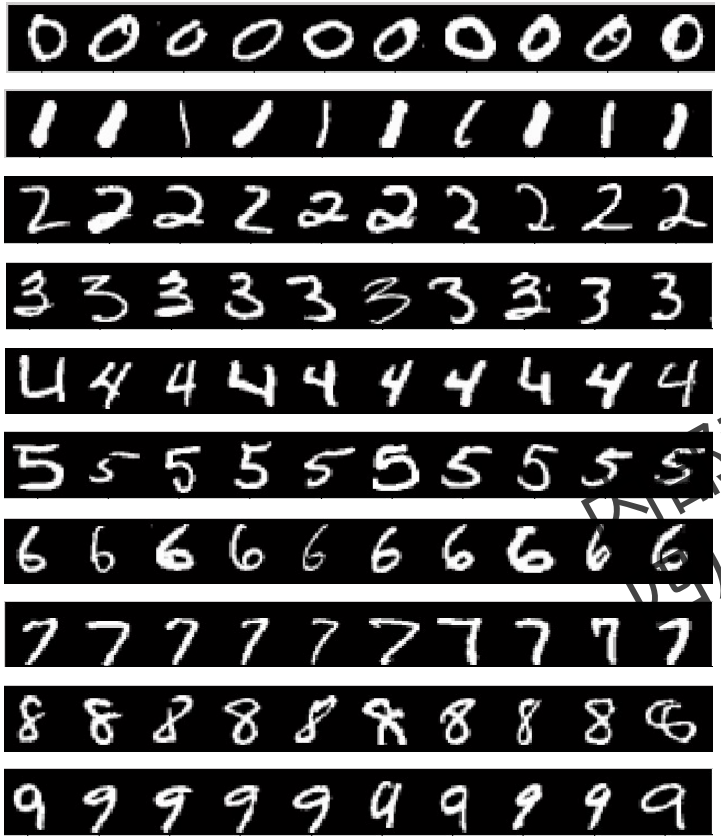


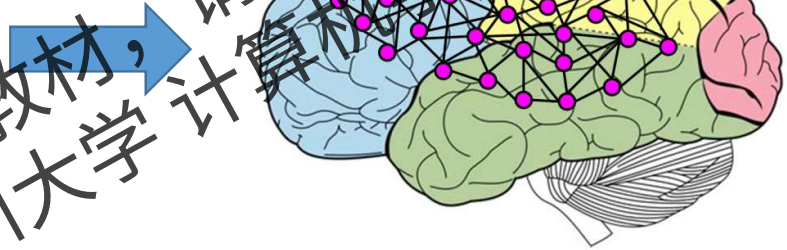


Introduction to brain structure



Introduction to brain structure



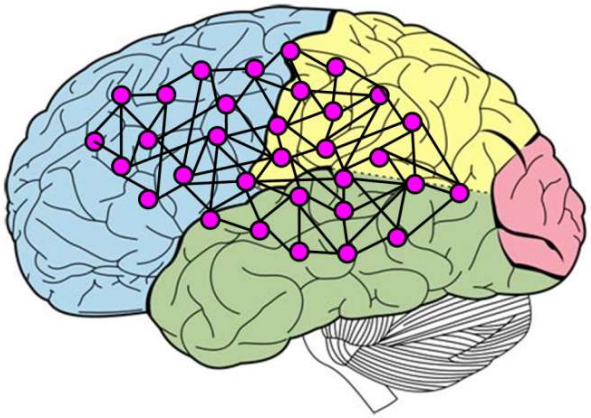


0
1
2
3
4
5
6
7
8
9

Problem:

- How to construct the computational model of brain?
- How to train the model of brain? How to develop the learning algorithm?

Introduction to brain structure



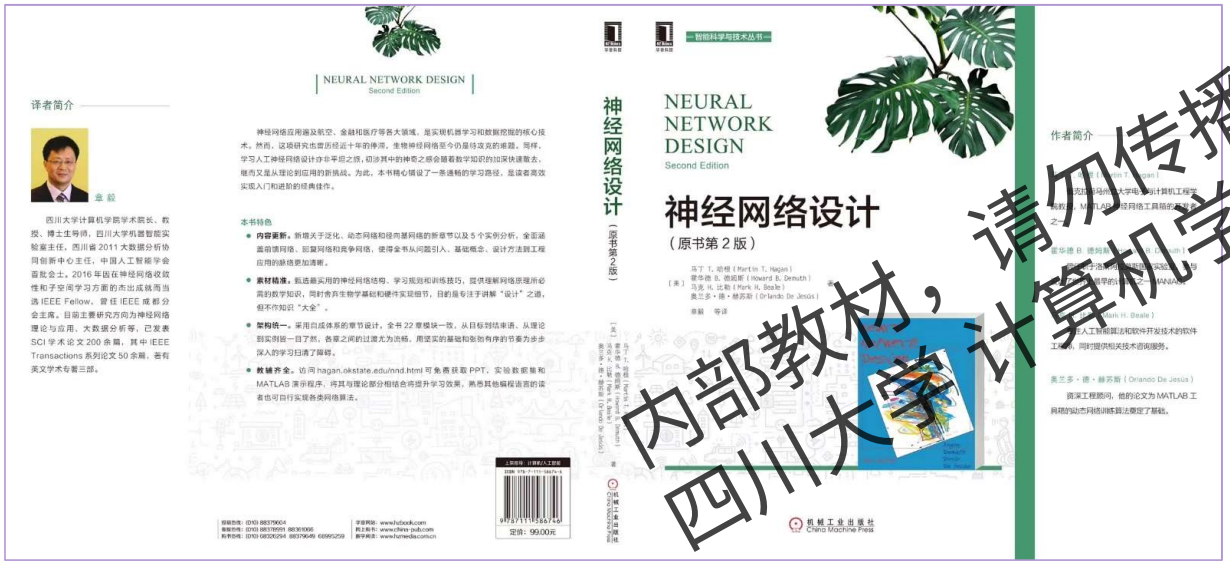
- What's next?
  - Constructing the network model.
  - Developing learning algorithm.

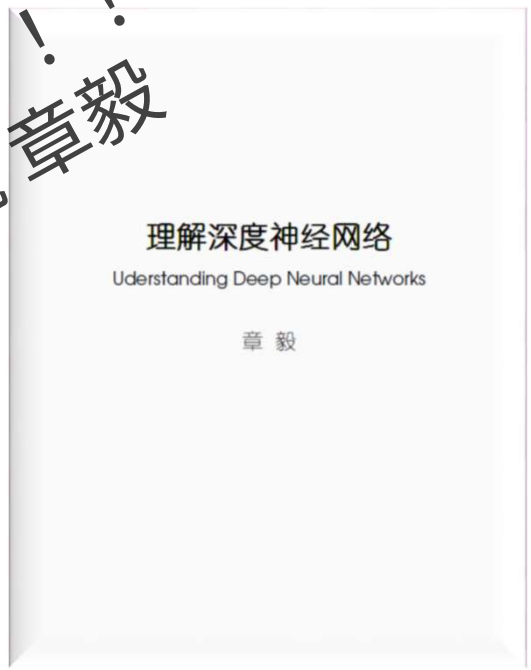
Outline

- Concepts
- An example: handwritten digits recognition
- How does a child recognize the handwritten digits
- Introduction to brain structure
- Discussions

Discussions

Reference Books





Under editing……

Discussions

TAs

Junjie Hu

Xiuyuan Xu

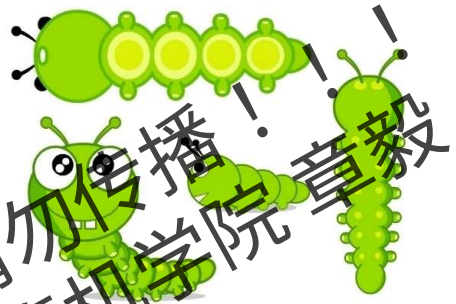
深度学习引论2018
扫一扫二维码，加入该群。

http://www.machineilab.org/

http://www.machineilab.org/users/zhangyi/index.html

Assignment

Implement the XOR-Worm problem classification by using MATLAB.



Doted worms

Smooth worms

$$F\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = f[f(2x_1 + 2x_2 - 1) + f(-x_1 - x_2 + 1.5) - 1.5]$$

$$f(s) = \begin{cases}1, & s \geq 0\\0, & otherwise\end{cases}$$

Thanks

内部教材，请勿传播！！！
四川大学 计算机学院 章毅