



厦门大学信息学院 本科选修课

2021-2022 第二学期

模式识别

Pattern Recognition

主讲：王程、陈龙彪



第六章

深度学习基础

陈龙彪

longbiaochen@xmu.edu.cn

厦门大学信息学院

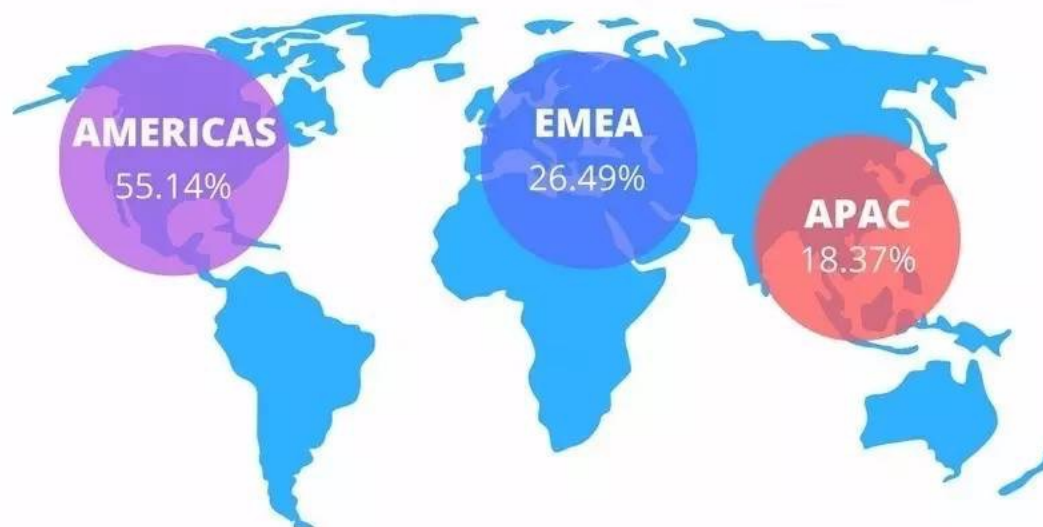
人工智能学习思维导图



深度学习得到大量关注

GLOBAL DEEP LEARNING SYSTEM MARKET 2016-2020 GEOGRAPHICAL SEGMENTATION

Global Deep Learning System Market by Geography 2015 (% share)



The market in the Americas is driven by the rapid increase in data generation across BFSI industries.

More than 1,500 companies are using AI in North America.

Since 2010, four of the biggest AI start-up acquisitions have happened in the UK.

APAC will be the fastest-growing segment, growing at a 41.58% CAGR through 2020.



新智元

机器学习 \approx 寻找一个函数

- 语音识别

$$f(\text{语音波形}) = \text{“How are you”}$$

- 图片识别

$$f(\text{小猫图片}) = \text{“Cat”}$$

- 下围棋

$$f(\text{围棋棋盘}) = \text{“5-5” (下一步棋的结果)}$$

- 对话系统

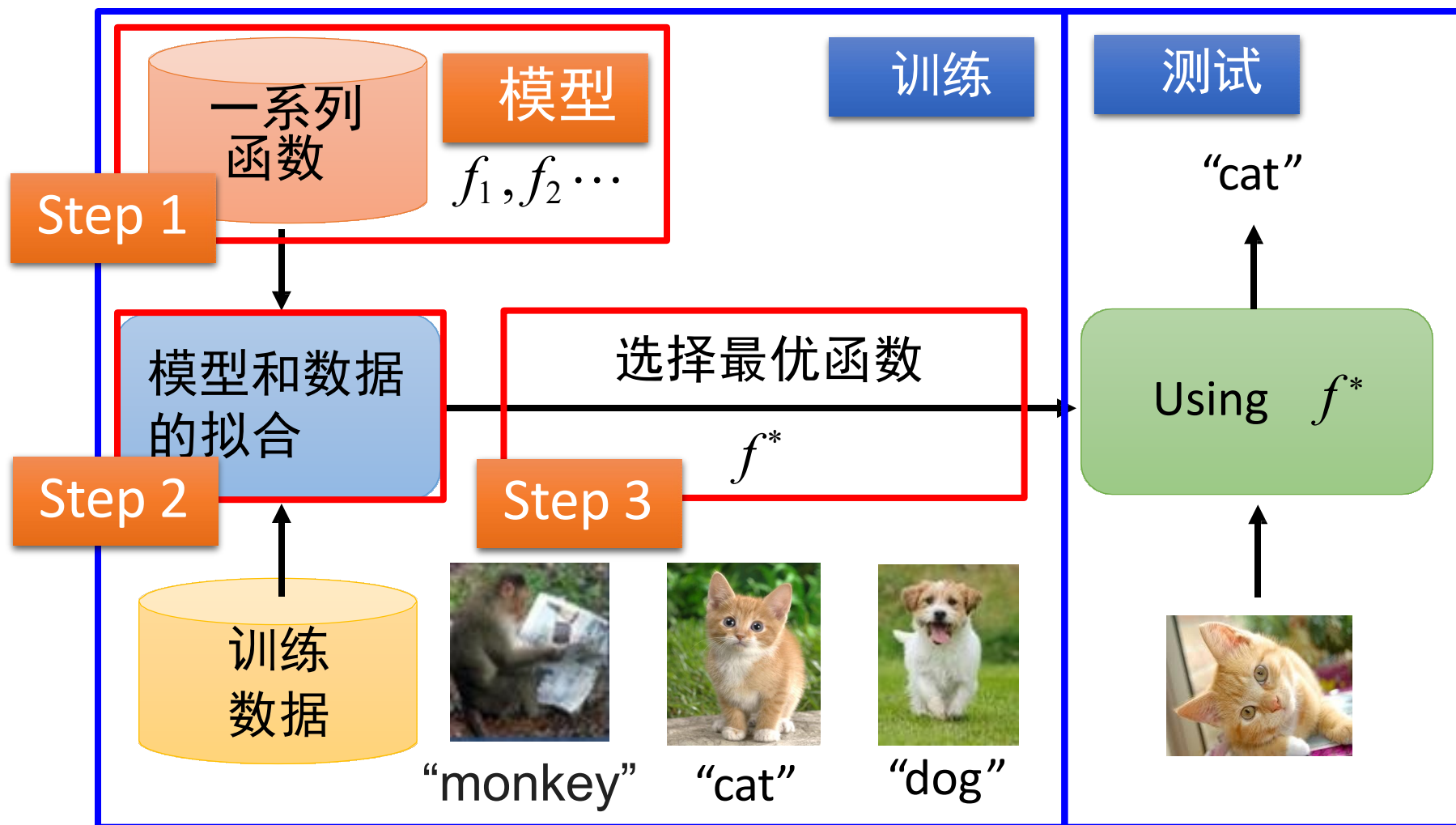
$$f(\text{“导航回家”}) = \text{“你说啥？”}$$

(你所说的内容) (系统的回答)

机器学习框架

图像识别:

$$f(\text{image of a cat}) = \text{"cat"}$$



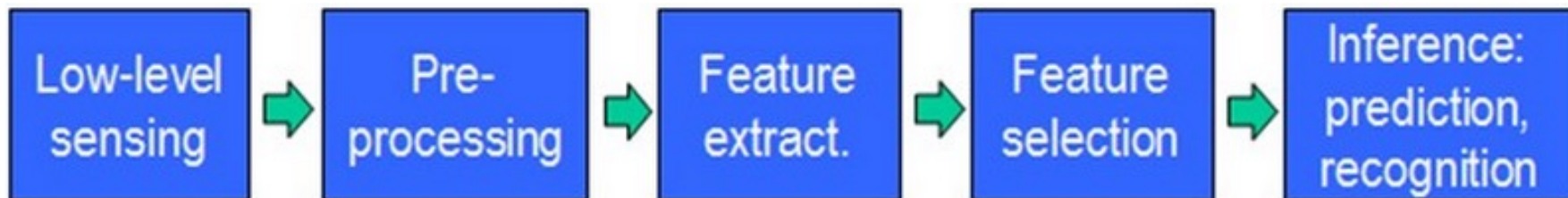
机器学习

- 机器学习（Machine Learning）是一门专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能的学科
- 简单地说，机器学习就是通过**算法**，使得机器能从大量的历史**数据**中学习规律，从而对新的样本做智能**识别或预测**未来

特征的自学习

□ 传统的模式识别方法

- 通过传感器获取数据，经过预处理、特征提取、特征选择、再到推理、预测或识别

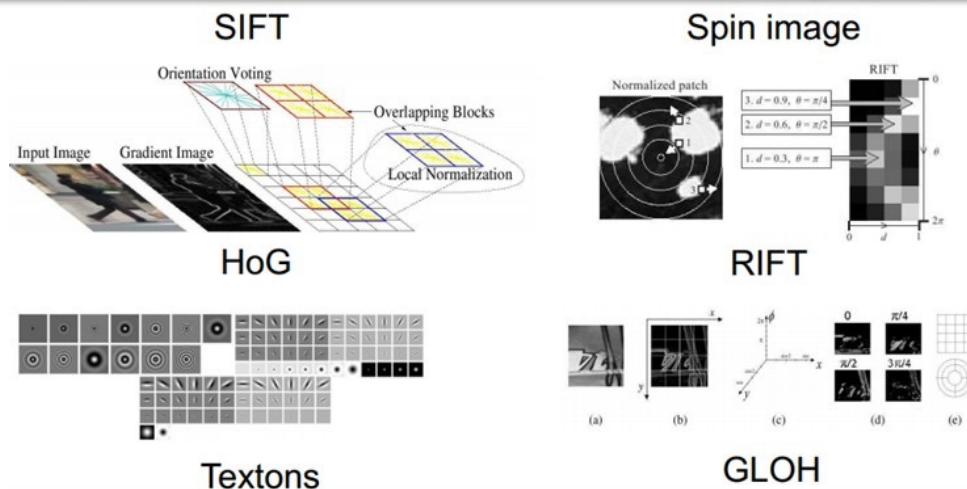


特征的自学习

□ 传统的模式识别方法

- 特征提取与选择的好坏对最终算法的确定性起关键作用
- 特征的样式一般都是靠人工提取特征，而手工选取特征费时费力，需要专业知识，很大程度上靠经验和运气

深度学习的出现为这个问题提出了一种解决方案

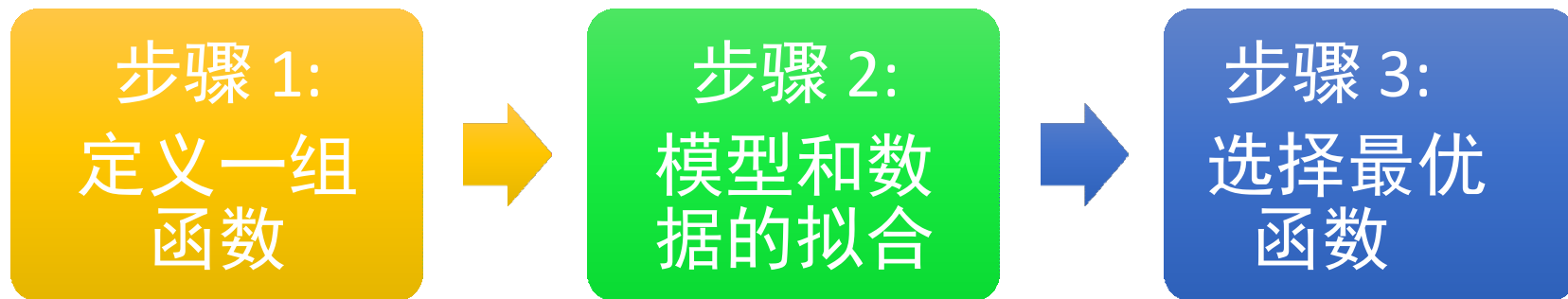


深度学习

- 自2006年，深度学习（Deep Learning）已经成为机器学习研究中的一个新兴领域。其动机在于建立、**模拟人脑**进行分析学习的神经网络，通过模拟人脑的机制来解释数据，例如图像，声音和文本
- 深度学习是机器学习的一种。深度学习的概念源于人工神经网络的研究，含多隐层的**多层感知器**就是一种深度学习结构。深度学习通过**组合低层特征形成更加抽象的高层表示**属性类别或特征，以发现数据的分布式特征表示

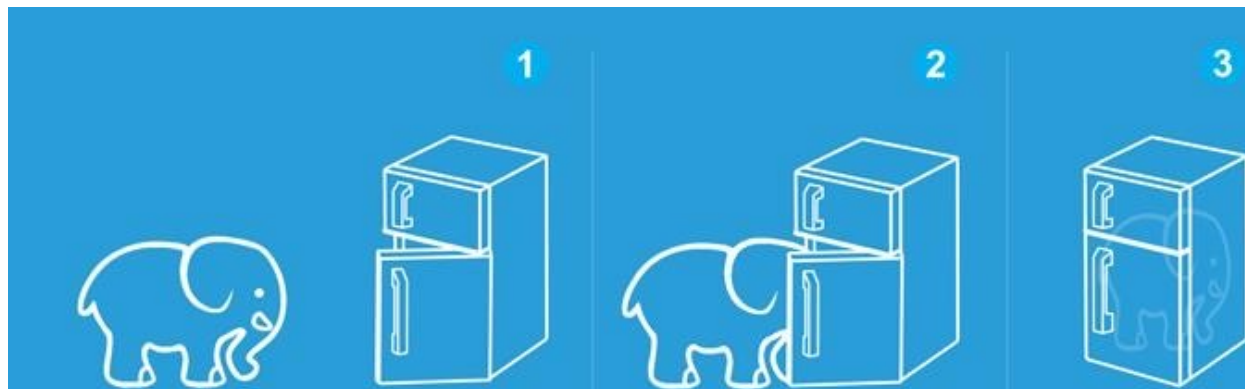
深度学习的三个步骤

□ 深度学习是如此简单...



问曰：把大象装进冰箱需要几步？

答曰：三步即可
把冰箱门打开，
把大象塞进去，
把冰箱门关上。

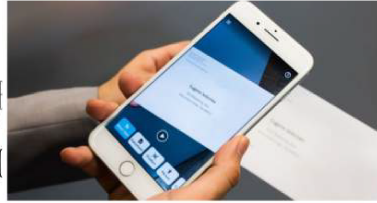


The Rise of Deep Learning

'Deep Voice' Software Can Clone Anyone's Voice With Just 3.7 Seconds of Audio

Using snippets of voices, Baidu's 'Deep Voice' can generate new speech, accents, and tones.

Let There Be Sight: How Deep Learning Is Helping the Blind 'See'



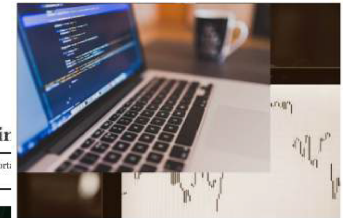
Technology outpacing security measures

Facial Recognition | Features and Interviews

AI beats docs in cancer spotting

A new study provides a fresh example of machine learning as an important diagnostic tool. Paul Riegler reports.

AI Can Help In Predicting Cryptocurrency Value



DEEPMIND I STARCRRAFT TRIUMPH FO

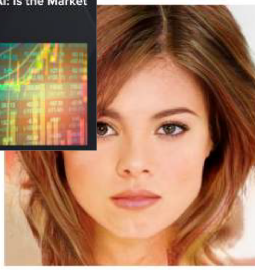
'Creative' AlphaZero leads way for chess computers and, maybe, science

Former chess world champion Garry Kasparov likes what he sees of computer that could be used to find cures for diseases

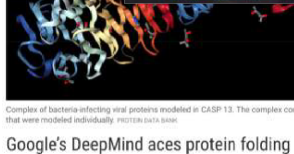


How an A.I. 'Cat-and-Mouse Game' Generates Believable Fake Photos

By CADE METZ and KEITH COLLINS JUN 2, 2018

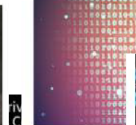


Stock Predictions Based On AI: Is the Market Truly Predictable?



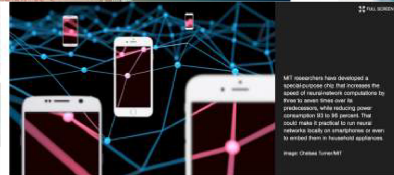
Google's DeepMind acers protein folding

By Robert F. Service | Dec. 4, 2018, 12:05 PM



On, Faked Data

18 JANUARY 24, 2018



Neural networks everywhere

New chip reduces neural networks' power consumption by up to 95 percent, making them practical for battery-powered devices.

Deep L

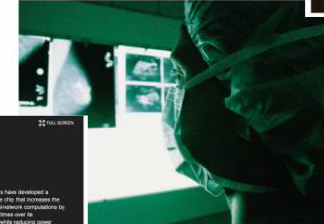
What, you'd/stop - Boston | Comment by Kenny Walker - Digital Reporter - @RandMagazine

After Millions of Trials, These Simulated Humans Learned to Do Perfect Backflips and Cartwheels

George Dornay 2018/12/26 - Physics 40+



Researchers introduce a deep learning method that converts mono audio recordings into 3D sounds using video scenes



3 faces show how far AI image generation has come in just four years

It's on the right aren't real: they're the product of machine learning



Automation And Algorithms: De-Risking Manufacturing With Artificial Intelligence

Sarah Goehrke Contributor @ Manufacturing | Focus on the industrialization of additive manufacturing.

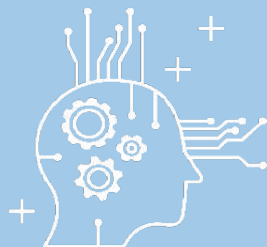
TWEET THIS

The two key applications of AI in manufacturing are pricing and manufacturability feedback

What is Deep Learning?

ARTIFICIAL INTELLIGENCE

Any technique that enables computers to mimic human behavior



MACHINE LEARNING

Ability to learn without explicitly being programmed



DEEP LEARNING

Extract patterns from data using neural networks

3 1 3 4 7 2
1 7 4 2 3 5

Learning representation

ML/AI: how do we learn features?

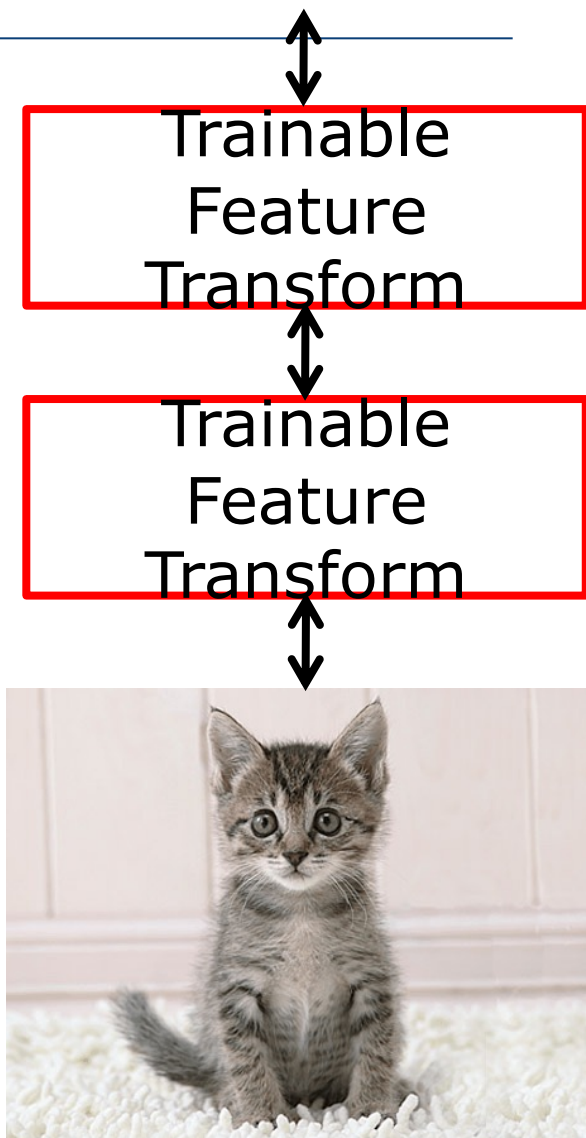
What is the fundamental rule?

What is the learning algorithms?

Neuroscience: how does the cortex learn perception?

CogSci: how does the mind learn abstract concepts on top of less abstract ones?

Deep learning addresses the problem of learning hierarchical representation with a single algorithm.

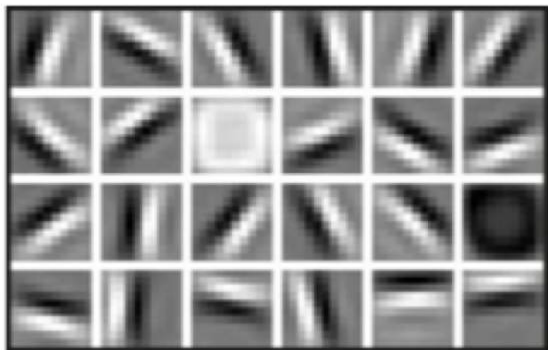


Why Deep Learning?

Hand engineered features are time consuming, brittle and not scalable in practice

Can we learn the **underlying features** directly from data?

Low Level Features



Lines & Edges

Mid Level Features



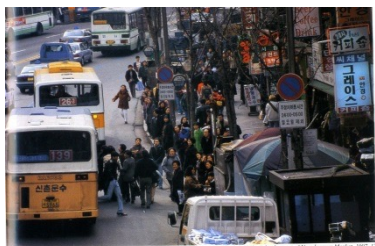
Eyes & Nose & Ears

High Level Features



Facial Structure

Is it brand new?



Neural Nets

McCulloch & Pitt 1943

Perception

Rosenblatt 1958

RNN

Grossberg 1973

CNN

Fukushima 1979

RBM

Hinton 1999

DBN

Hinton 2006

D-AE

Vincent 2008

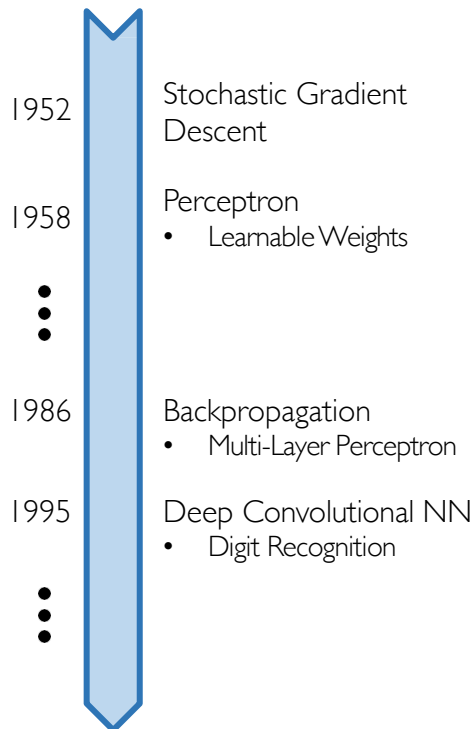
AlexNet

Alex 2012

GoogLeNet

Szegedy 2015

Why Now?



Neural Networks date back decades, so why the resurgence?

1. Big Data

- Larger Datasets
- Easier Collection & Storage

IMAGENET



2. Hardware

- Graphics Processing Units (GPUs)
- Massively Parallelizable

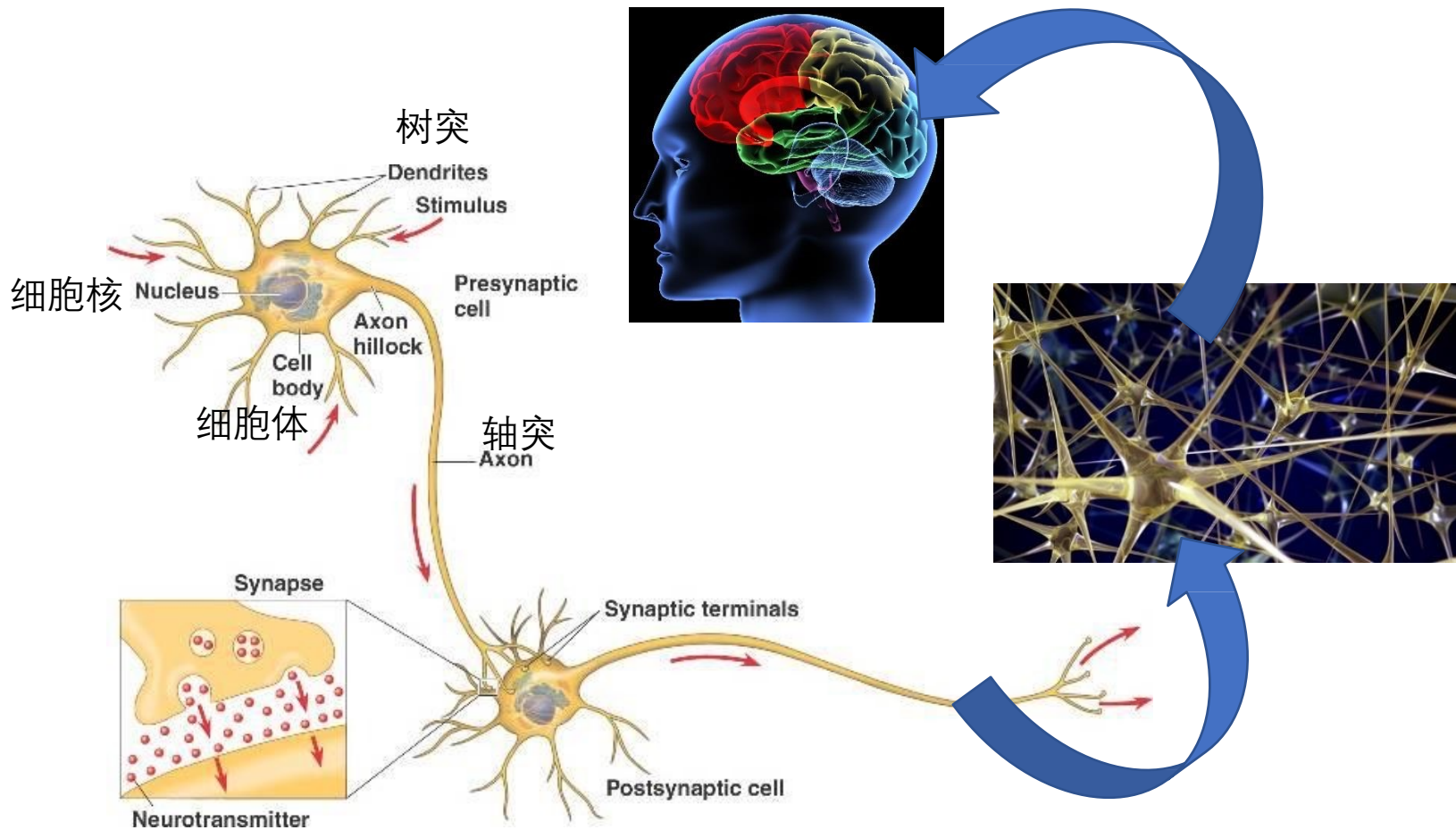


3. Software

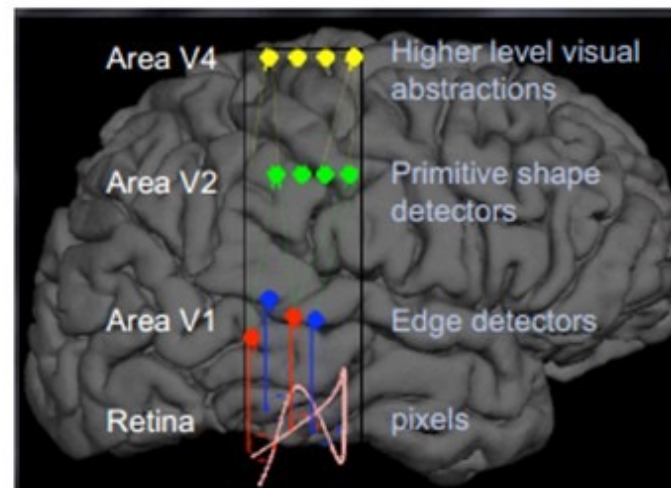
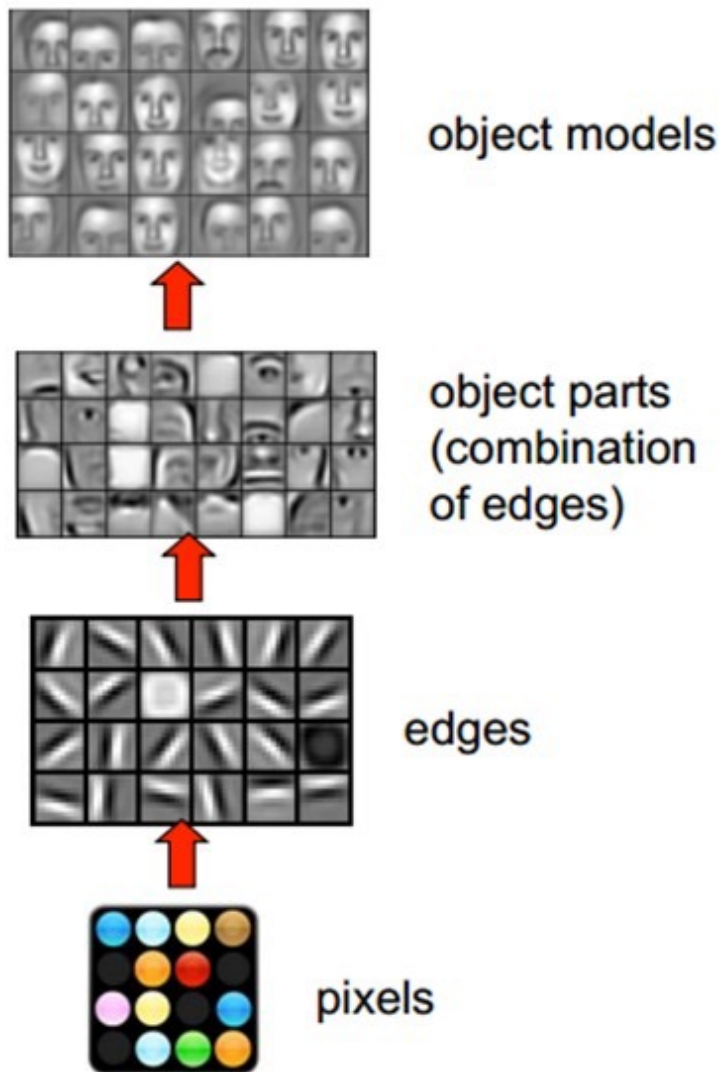
- Improved Techniques
- New Models
- Toolboxes



人类的大脑



人脑的视觉原理



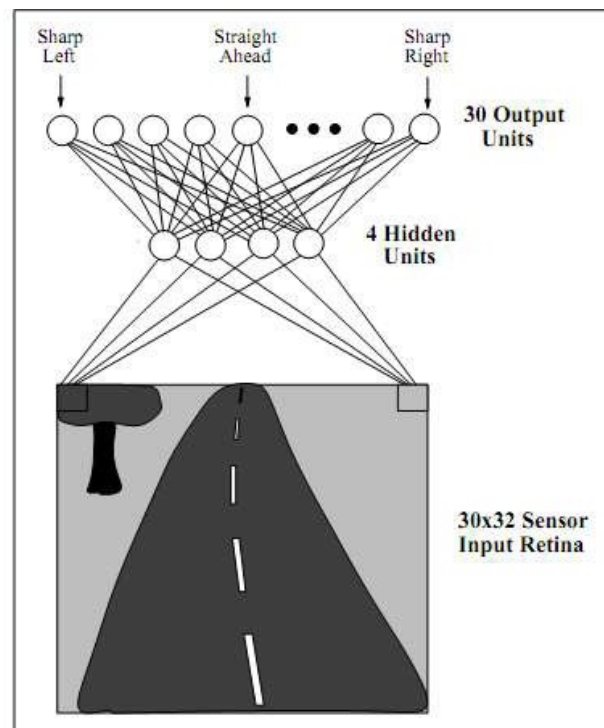
深度学习

- 受到大脑结构分层的启发，神经网络的研究发现多隐层的人工神经网络具有优异的特征学习能力
- 学习得到的特征对数据有更本质的刻画，从而有利于可视化或分类
- 深度学习可以通过学习一种深层非线性网络结构，实现复杂函数逼近，表征输入数据分布式表示

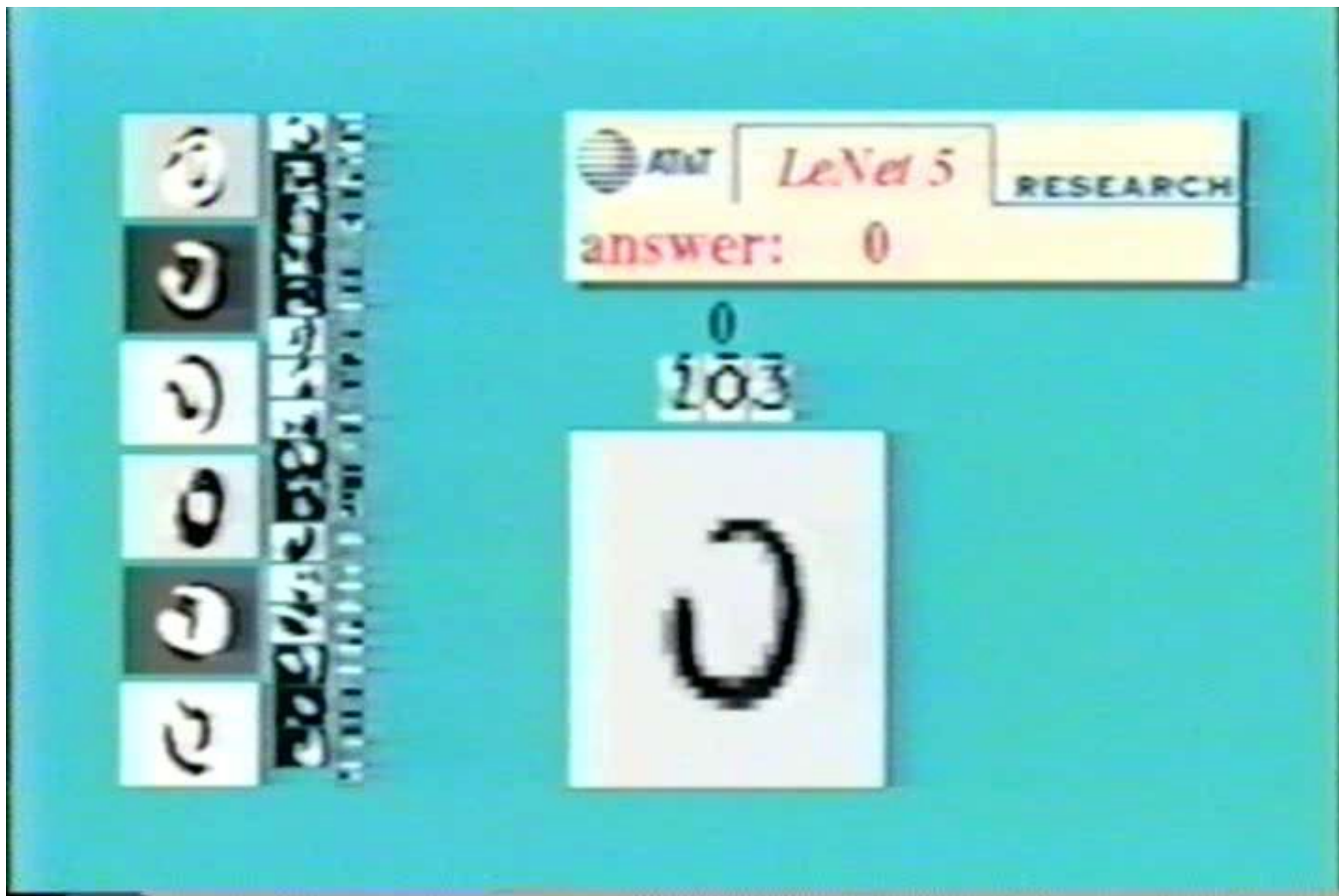
Applications

Neural network application

- ALVINN drives 70 mph on highways



Neural network application



Scene Recognition (CNN)



Predictions:

- **Type of environment:** outdoor
- **Semantic categories:** rock_arch:0.63, arch:0.30,
- **SUN scene attributes:** rugged, natural light, dry, climbing, far-away horizon, touring, rocky, open area, warm, sand

B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, and A. Oliva. "Learning Deep Features for Scene Recognition using Places Database." Advances in Neural Information Processing Systems 27 (NIPS), 2014.

Visual Style Recognition (CNN)



HDR



Macro



Baroque



Rococo



Vintage



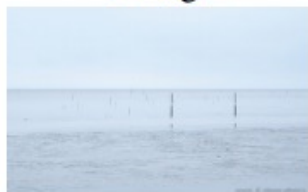
Noir



Northern Renaissance



Cubism



Minimal



Hazy



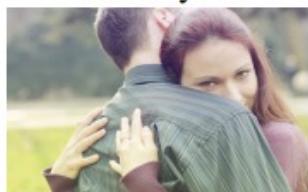
Impressionism



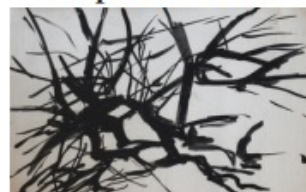
Post-Impressionism



Long Exposure



Romantic

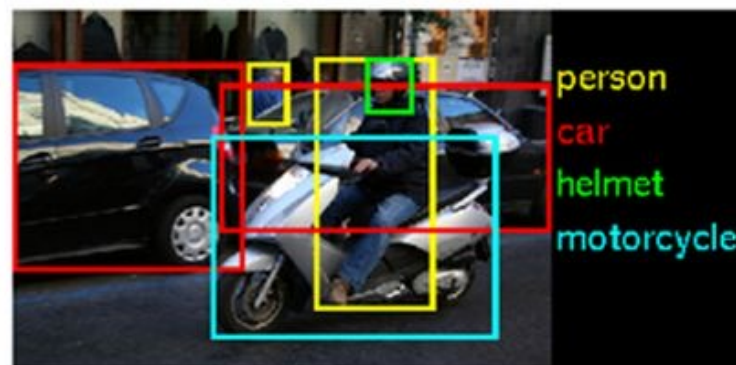
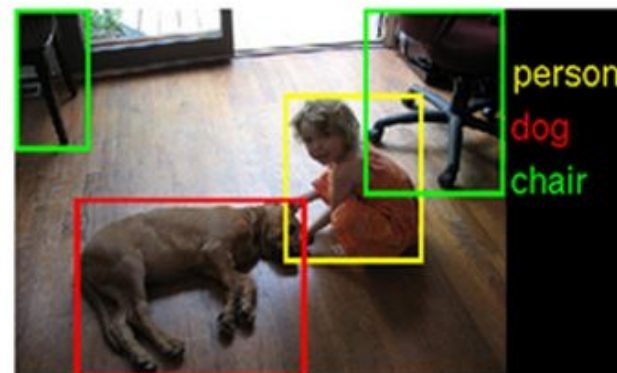
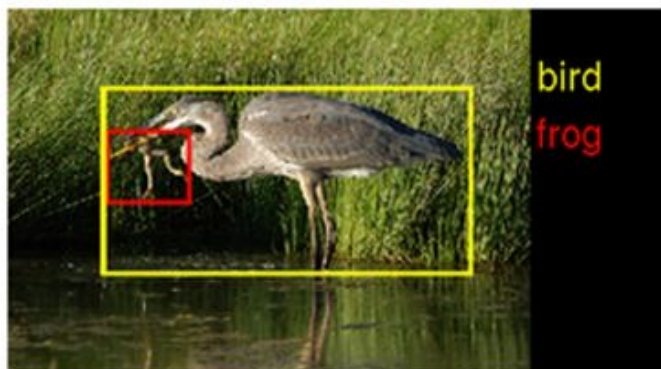


Abs. Expressionism



Color Field Painting

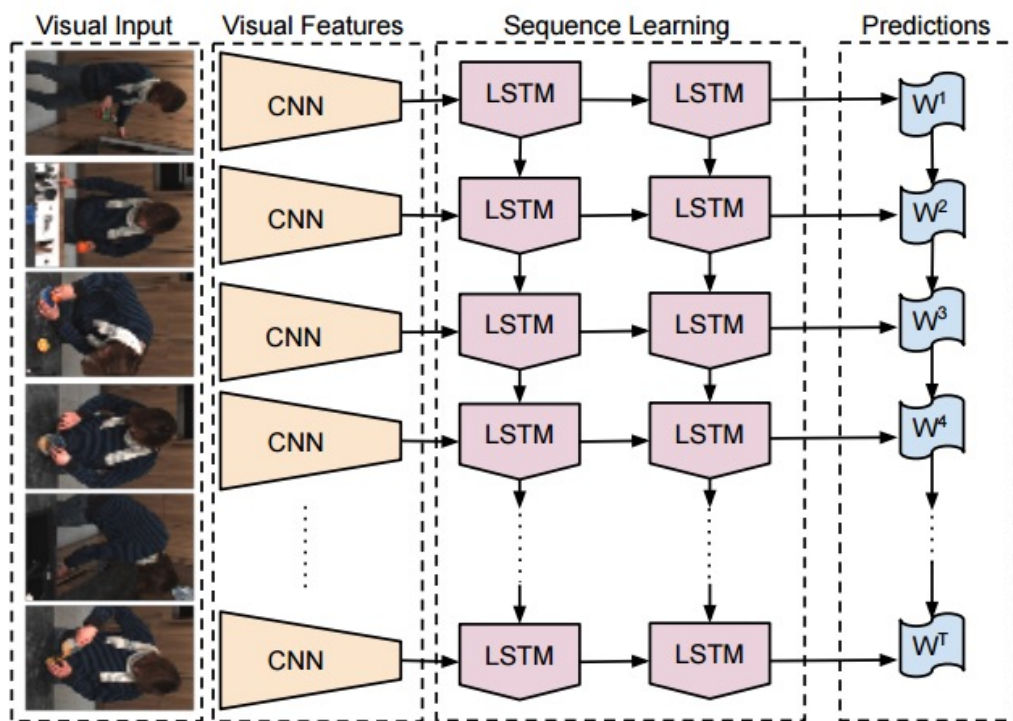
Object Detection (R-CNN)



Detection \approx Localization + Classification

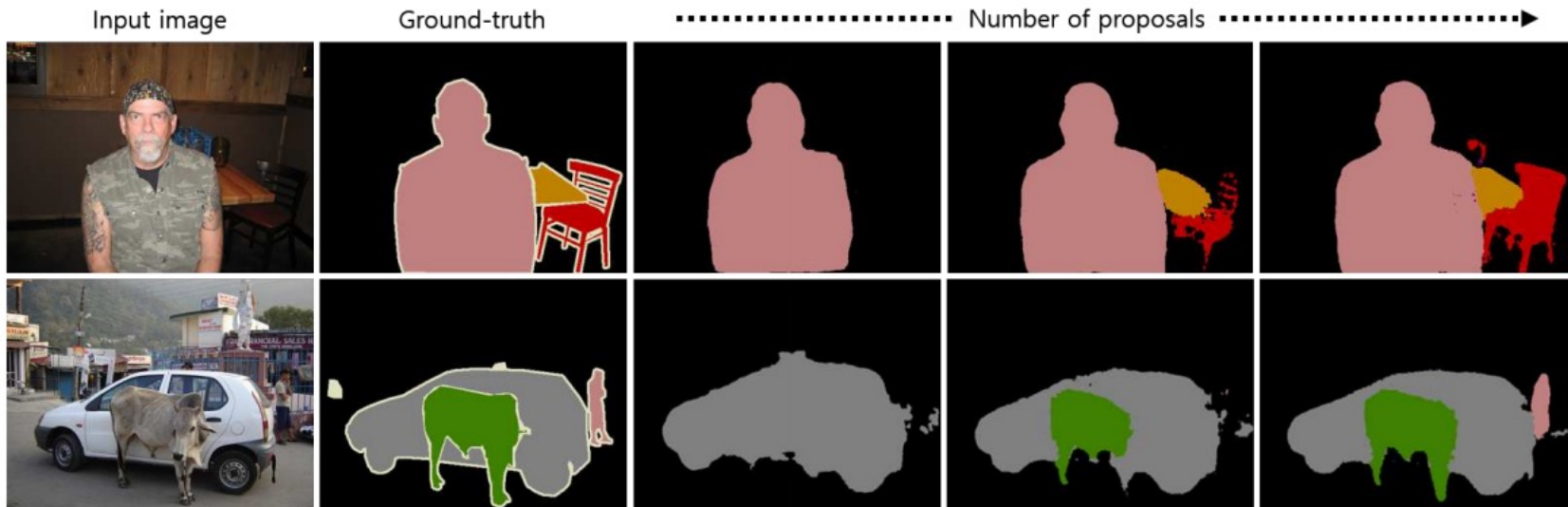
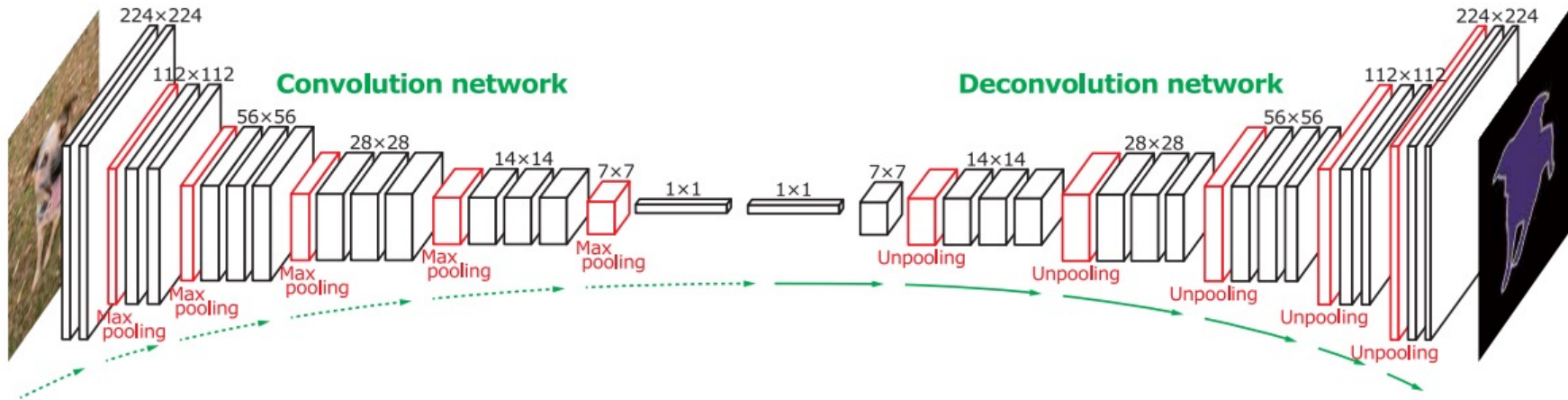
Girshick, Ross, et al. "Region-based convolutional networks for accurate object detection and segmentation." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 38.1 (2016): 142-158.

Image Captioning (CNN+LSTM)



A black and white cat is sitting on a chair.

Segmentation (DeconvNet)



Deep Visuomotor Control (CNN)

Learned Visual Representations

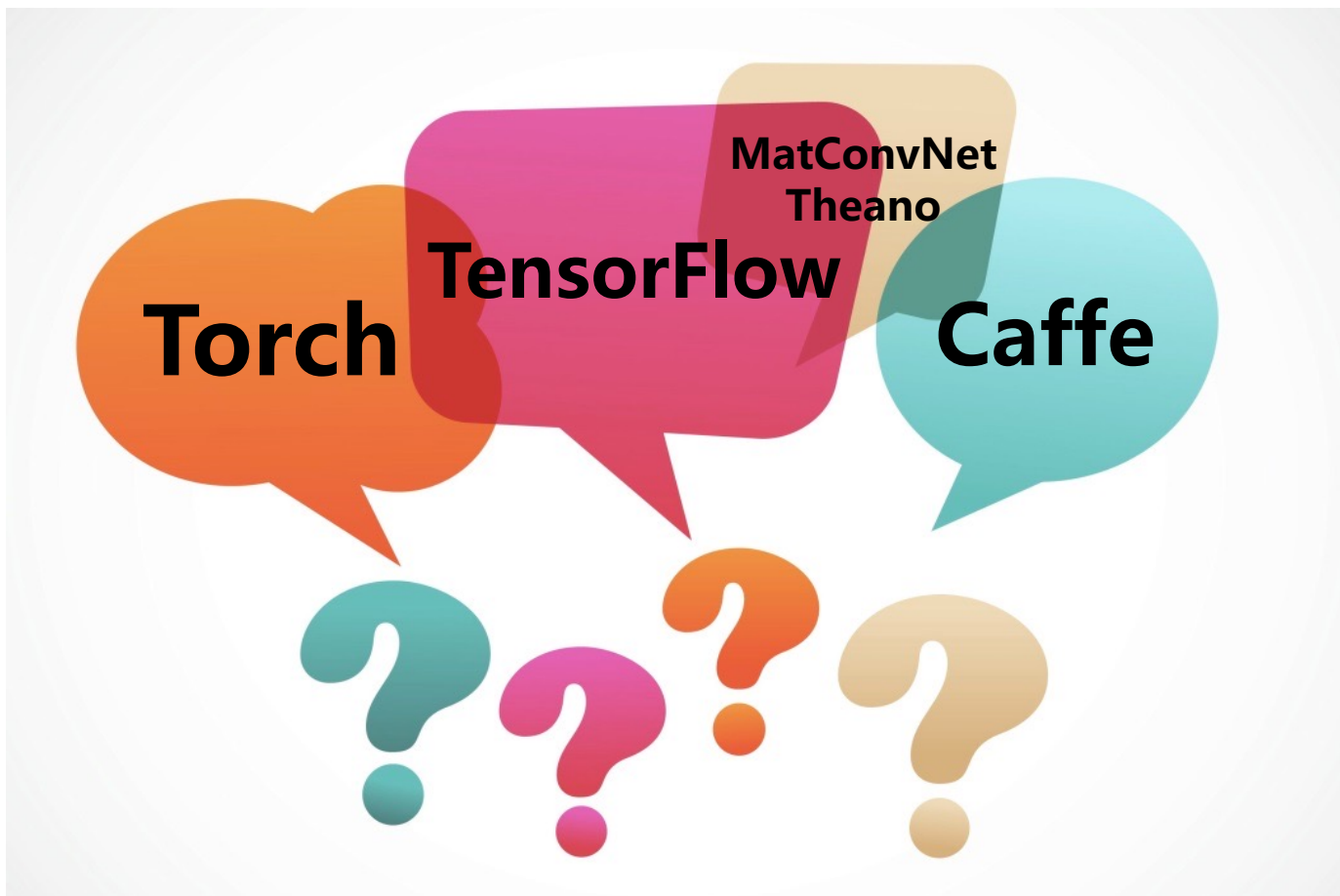
Neural Style (CNN)



Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "A neural algorithm of artistic style." arXiv, 2015

Tools

Deep learning tools?



Caffe

From U. C. Berkeley

Written in C++

Good for training or fine-tuning feedforward models

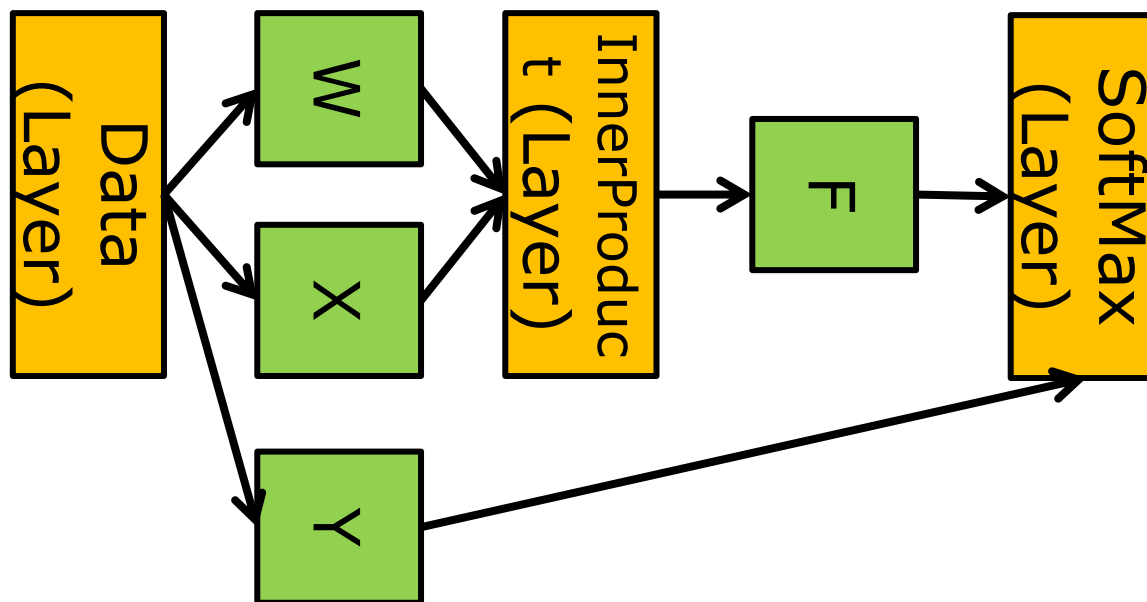
Caffe

Blobs: Store data and derivatives

Layer: Transform bottom blobs and top blobs

Net: Many layers; computes gradient via forward / backward

Solver: Uses gradients to update weights



Caffe

- (+) Good for feedforward networks (CNN)
- (+) Good for fine-tuning existing networks
- (+) Train models without writing any code
- (+) Python interface is pretty useful
- (-) Need to write C++ / CUDA for new GPU layers
- (-) Not good for recurrent networks
- (-) Cumbersome for big networks (GoogLeNet, ResNet)

Torch

From NYU

Written in C and Lua

Used a lot in Facebook and DeepMind

Torch

- (+) Easy to write your own layer types and run on GPU
- (+) Many available library codes and packages
- (+) Lots of pretrained models
- (+) Lots of modular pieces that you can combine
- (-) Lua
- (-) Less plug-and-play than Caffe
- (-) Not great for RNNs

Theano

From University of Montreal

Embracing computation graphs, symbolic computation

High-level wrappers: Keras, Lasagne

Theano

- (+) Python + numpy
- (+) Computational graph is cool
- (+) Good for RNNs
- (+) High level wrappers ease the pain
- (-) Not enough pretrained models
- (-) Error messages can be unhelpful
- (-) Raw Theano is somewhat low-level

TensorFlow

From Google

Very similar to Theano

Easy visualizations

Multi-GPU and multi-node training

TensorFlow

- (+) Python + numpy
- (+) Computational graph is cool
- (+) Much faster than Theano
- (+) Easy to modify nets
- (-) Not many pretrained model (but it will get better)
- (-) Training time is slower than others right now
- (-) Can hack into the networks

White to black



Overview

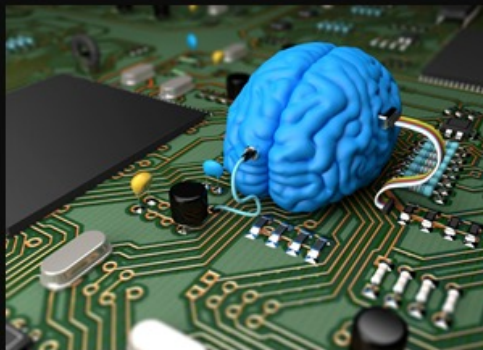
	Caffe	Torch	Theano	TensorFlow
Language	C++, Python	Lua	Python	Python
Pretrained	Yes ++	Yes ++	Yes (Lasagne)	Inception
Multi-GPU: Data parallel	Yes	Yes <small>cunn. DataParallelTable</small>	Yes <small>platoon</small>	Yes
Multi-GPU: Model parallel	No	Yes <small>fbcunn.ModelParallel</small>	Experimental	Yes (best)
Readable source code	Yes (C++)	Yes (Lua)	No	No
Good at RNN	No	Mediocre	Yes	Yes (best)

训练方法

Deep Learning



What society thinks I do



What my friends think I do



What other computer scientists think I do



What mathematicians think I do



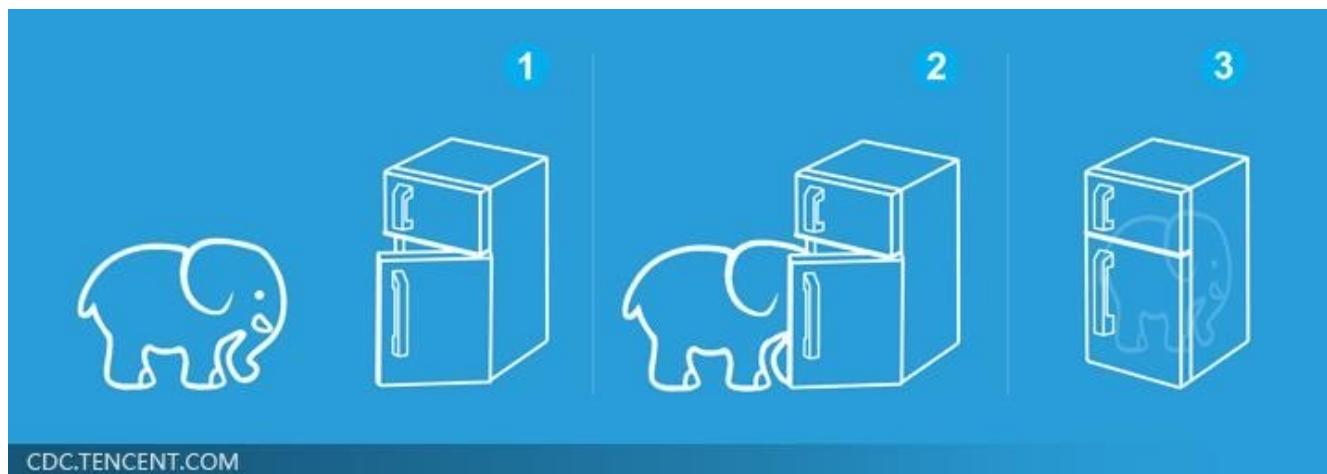
What I think I do

```
from theano import *
```

What I actually do

结束语

□ 深度学习是如此简单...



深度学习

□ 深度学习的实质

- 通过构建具有**很多隐层的机器学习模型**，在**海量的训练数据上学习有用的特征**，最终提升**分类或预测的准确性**
- **深度模型**是手段，**特征学习**是目的

深度学习Demo

□ 手写字符识别

- <https://transcranial.github.io/keras-js/#/mnist-cnn>

□ 图像识别

- [黑天鹅](#)

□ 超像素还原

- <https://transcranial.github.io/keras-js/#/image-super-resolution>

参考资料

□ Hung-yi Lee, Machine Learning, 2019

- <http://speech.ee.ntu.edu.tw/~tlkagk/courses.html>

□ dl_tutorials

- https://github.com/sjchoi86/dl_tutorials

□ Keras.js

- <https://github.com/transcranial/keras-js>



End

