



SeetaTech
中 科 视 拓

深度学习简史

主讲人：刘昕 博士

当谈到深度学习的时候，大家在说什么？



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



目录 CONTENTS

1

深度学习引导

2

深度学习简史

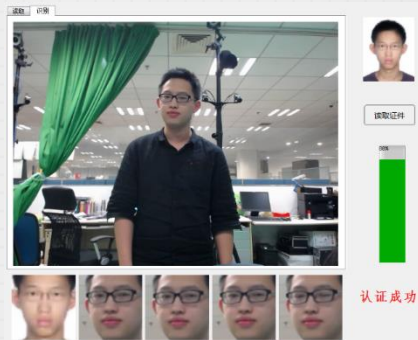
3

总结与展望



AI大热的背后：跨越式的进步

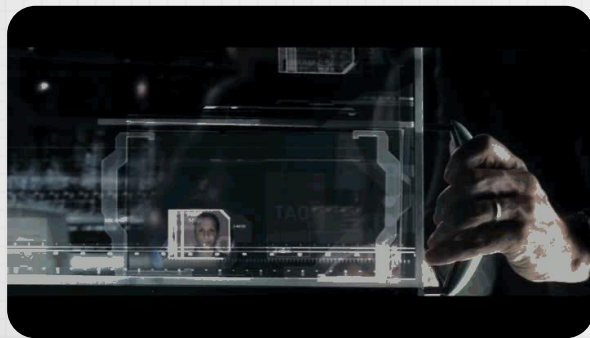
- AI正在被快速迭代发展！
 - 例1：惊破天的AlphaGo
 - 例2：汽车辅助驾驶&自动驾驶技术
 - 例3：超预期发展的人脸识别技术
 - 例4：智能音响（人机对话）
 - 例5：高考机器人



1968 星际迷航



2002 少数派报告



2010



正在实现！



深度学习三驾马车

- 深度模型（1943--）
 - 1943 MCP模型
 - 1957 单层感知机
 - 1969 Perception一书宣判死刑
 - 1986（1974）年误差反向传导
 - 2012 卷积神经网络（CNN）



- 大数据（互联网，传感器）
- 高性能计算（GPU）

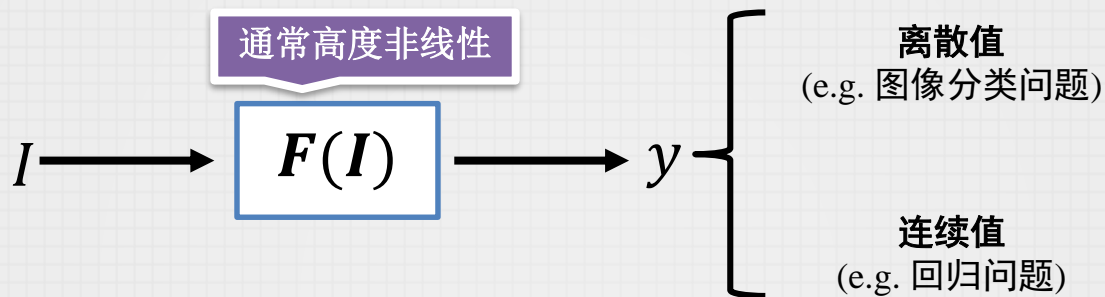


SeetaTech
中科视拓

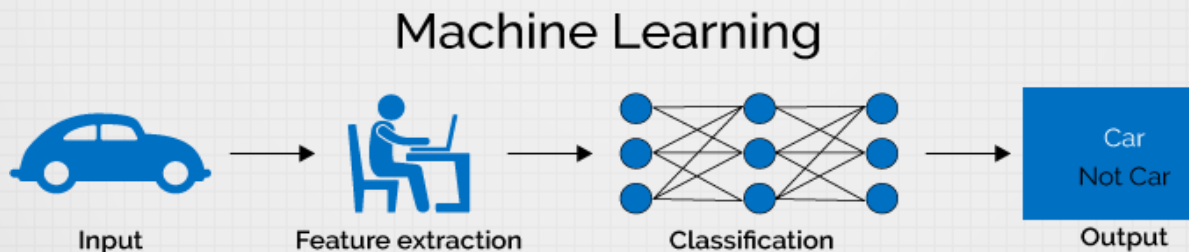


深度学习方法模型

➤ 方法模型



➤ 经典的两段式方法



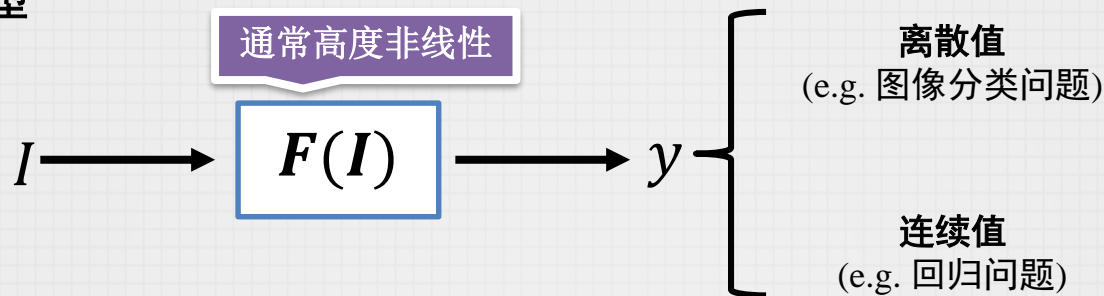
前深度学习学习方法模型：以计算机视觉为例

计算机视觉两段式方法的一个漫画



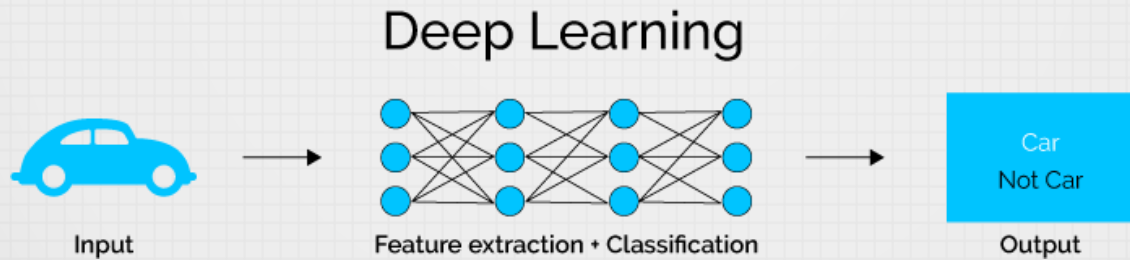
深度学习方法模型

方法模型



深度学习带来的方法革命

- ▣ 显示学习非线性映射 $F(I)$
- ▣ 端到端学习(End to End) / 协同增效



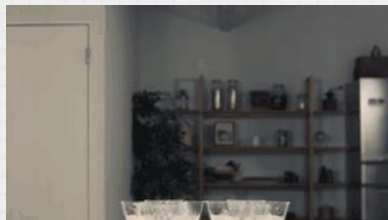
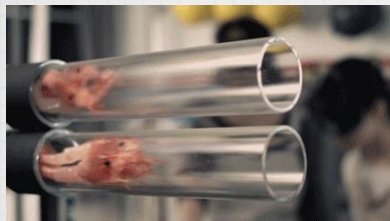
Traditional Cooking

➤ 人工拆分步骤、设计参数



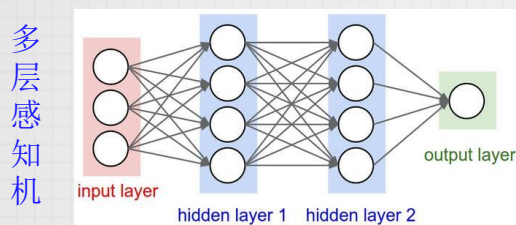
Deep Cooking

➤ 算法自动划分步骤、学习参数



深度学习是什么？

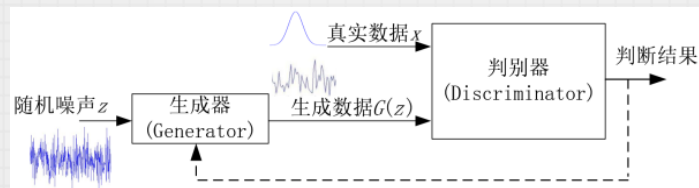
- 以不少于两个隐含层的神经网络对输入进行“端到端”的非线性变换或表示学习的技术
- 包括多种结构：MLP, CNN, RNN/LSTM, GAN等



卷积神经网络

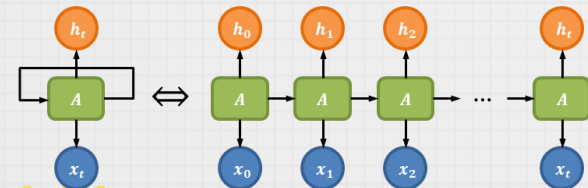


GAN:

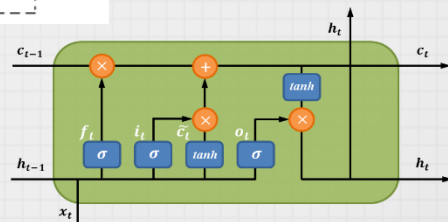


RNN

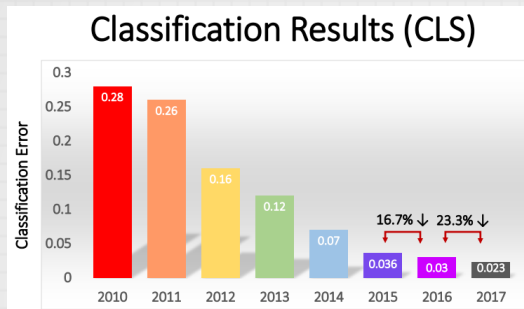
输出层
隐含层
输入层



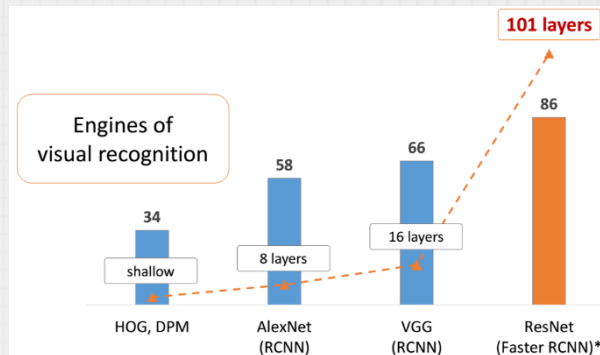
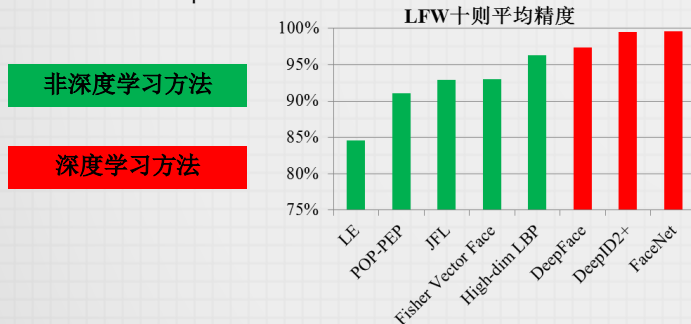
LSTM



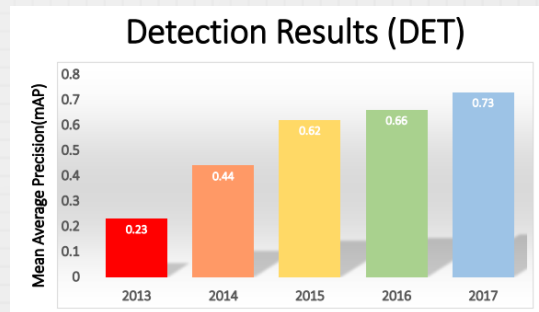
深度学习带来的CV领域的突破性进步



ImageNet图像分类Top-5错误率



Pascal VOC目标检测MAP



ImageNet目标检测MAP



深度学习还能做什么？

- **专家系统**：问答系统，决策系统（AlphaGo）
- **计算机视觉**：三维重建，生物特征识别，图像与视频理解...
- **语音识别**：语音输入，语音控制，远场语音识别（智能音响）
- **自然语言理解**：情感分析（游戏、舆情），聊天机器人
- **金融风控**：征信，风险评估，坏账预警
- **智能投顾**：程序化交易
- **智慧医疗**：机器读片，疾病预测，基因诊断

.....



深度学习背后的男人



左起:

LeCun (CNN)

Hinton (06深度学习大爆发)

Bengio (侠之大者)

吴恩达

目录 CONTENTS

1

深度学习引导

2

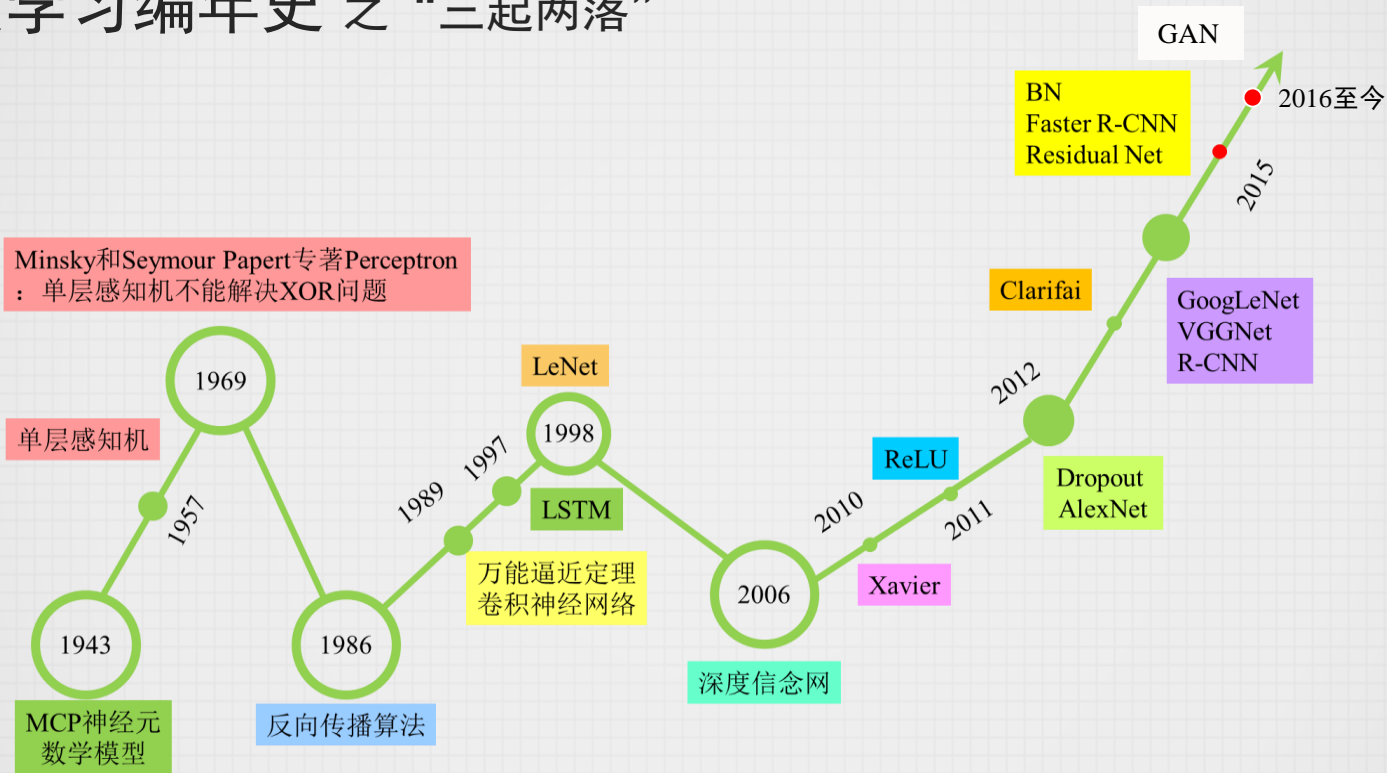
深度学习简史

3

总结与发展

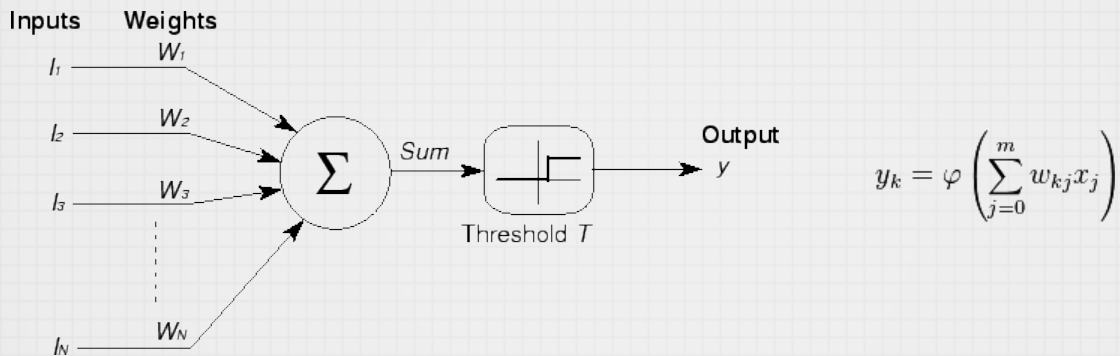


深度学习编年史 之 “三起两落”



一起：McCulloch-Pitts Model of Neuron 1943

➤ 单个神经元的计算模拟

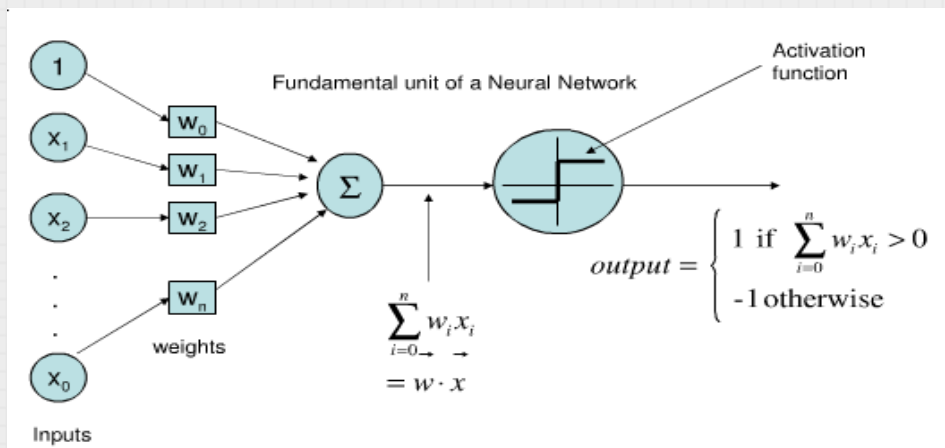


MCP人工神经元模型：多个input输入，权重，内积运算，激活函数，输出，神经元之间传递



一起：Rosenblatt 1958年感知机模型

➤ 最早用于二类分类问题



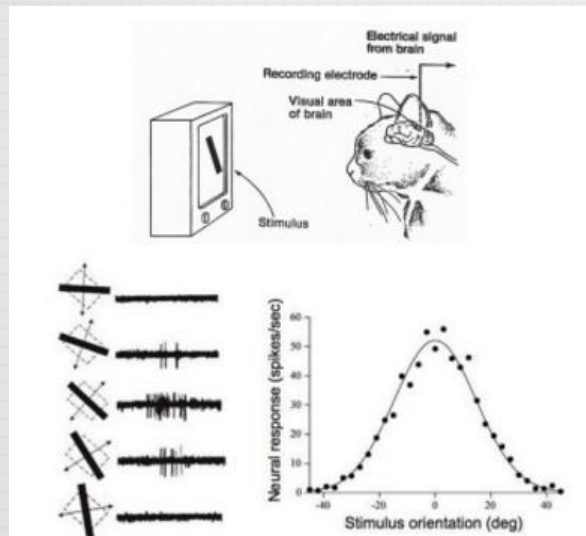
➤ 采用梯度下降方法求解



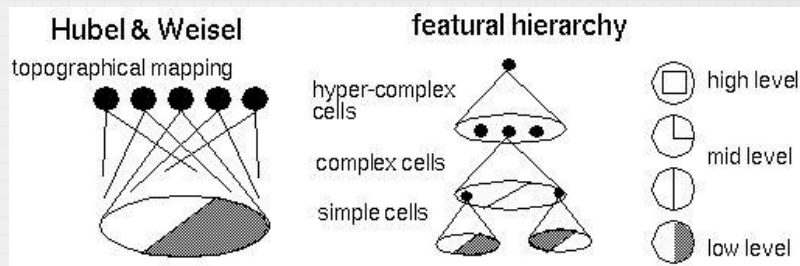
一起：Hubel & Wiesel 1962

- Hubel & Wiesel对视觉皮层（Visual Cortex）的功能划分：从简单特征提取神经元（简单细胞）到渐进复杂的特征提取神经元（复杂细胞，超复杂细胞等）的层级连接结构

研究方法：动物实验



模型：



功能划分：

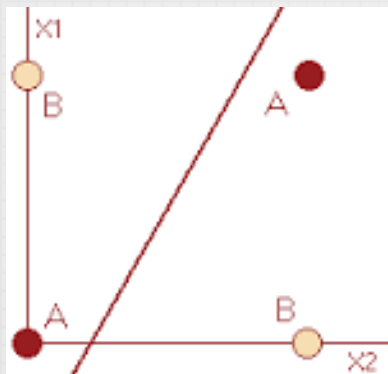
Cell Type	Selectivity
Simple	orientation, position
Complex	orientation, motion, direction
Hypercomplex	orientation, motion, direction, length

Receptive fields, binocular interaction and functional architecture in the cat's visual cortex



一落：单层感知机模型的缺陷

- 1962年，单层感知机模型的收敛性得到理论证明[1]
- Minsky & Papert的专著*Perceptron*(1969)指出单层感知机解决不了XOR(异或)问题
- 本质上是线性模型

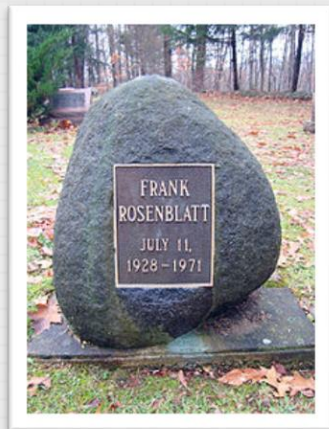


Novikoff, A. B. J. (1962). On convergence proofs on perceptrons. Proceedings of the Symposium on the Mathematical Theory of Automata (pp. 615–622)
XOR problem theory:
http://home.agh.edu.pl/~vlsi/AI/xor_t/en/main.htm



一落：单层感知机模型的缺陷

- 几乎宣判了单层感知机甚至人工神经网络的死刑，导致了后来NN研究多年的寒冬
- Rosenblatt 1971年43周岁生日时因游船事故去世，没有等到NN研究第二次春天的到来



IEEE Frank Rosenblatt Award

2014 - Geoffrey E. Hinton

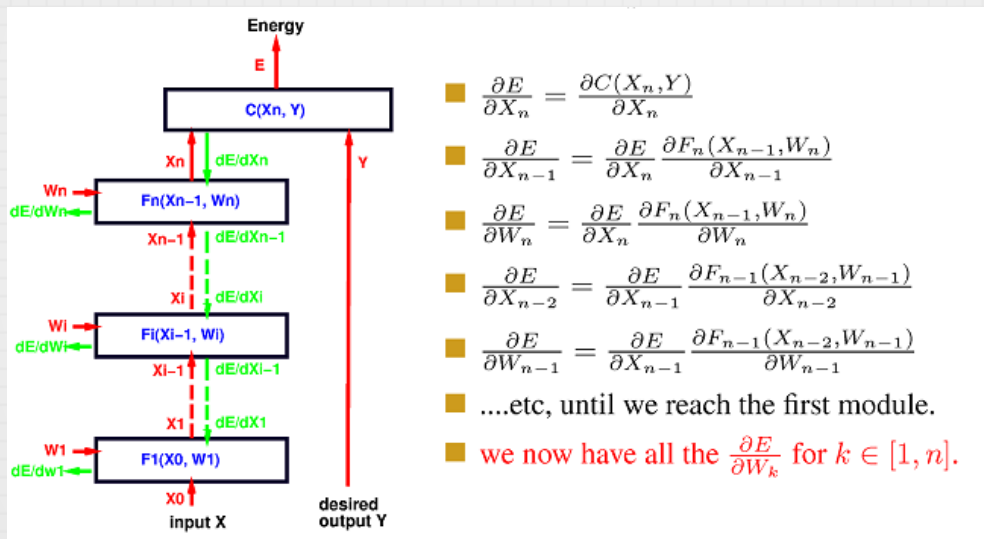
2012 - Vladimir N. Vapnik

2009 - John J. Hopfield



二起：Back Propagation

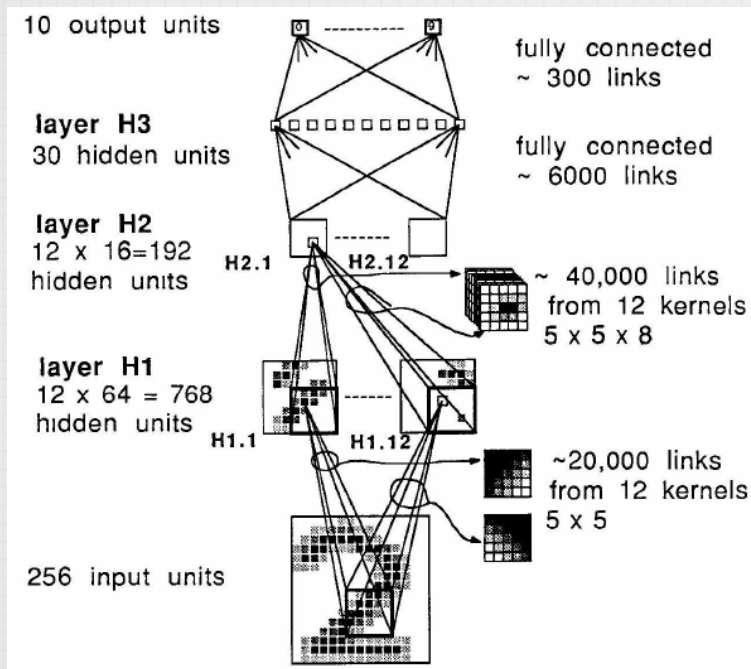
- 1974年Webos在博士论文中首次提出BP算法，但未引发关注
- 目前广泛使用的BP算法诞生于1986年
- 以全连接层为例：链式求导，梯度反向传导



Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (8 October 1986). "Learning representations by back-propagating errors". *Nature* **323** (6088): 533–536.



二起: First Baby: LeCun1989



Highlights:

- 1) 3个隐层, 已经可以称为是deep network
- 2) Response map之间也共享卷积核权重

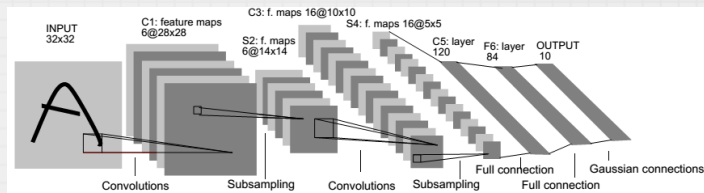
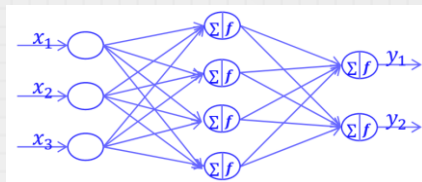
Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel: Backpropagation Applied to Handwritten Zip Code Recognition, *Neural Computation*, 1(4):541-551, Winter 1989



二落：1980s~2006

■ 成就

- **Error梯度反向传播**
- 多层感知机（MLP）
- 卷积网络（LeCun1989, LeNet）
- 非线性激活函数
- 万能逼近定理（1989）



■ 缺陷

- **优化困难：梯度消失**
- 训练数据不足
- 计算资源有限



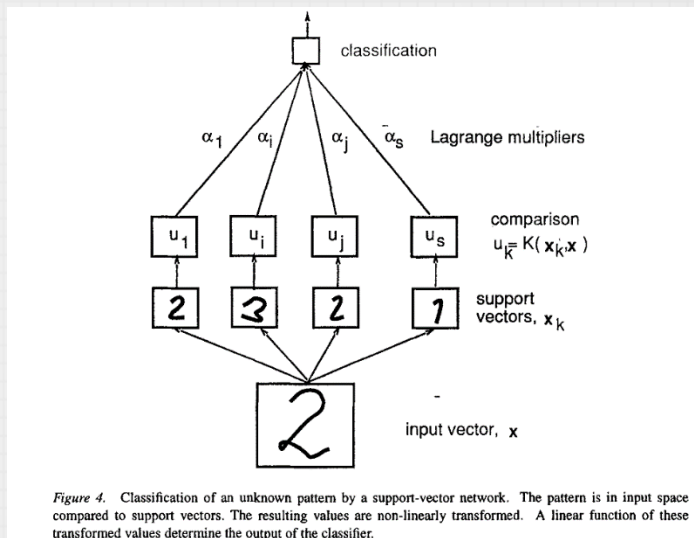
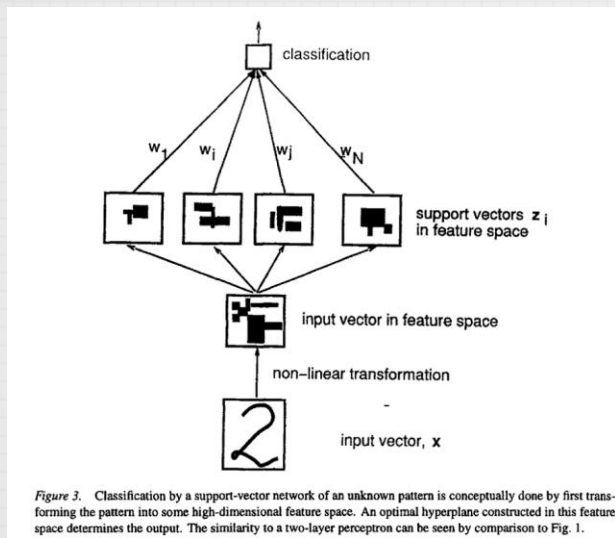
- Decision Tree
- SVM
- Boosting
- Sparse Coding
- Graph Model

Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (8 October 1986). "Learning representations by back-propagating errors". *Nature* **323** (6088): 533–536.



二落：SVM和NN不可不说的故事

➤ 95年Vapnik那篇著名文章的题目叫“Support Vector Network”

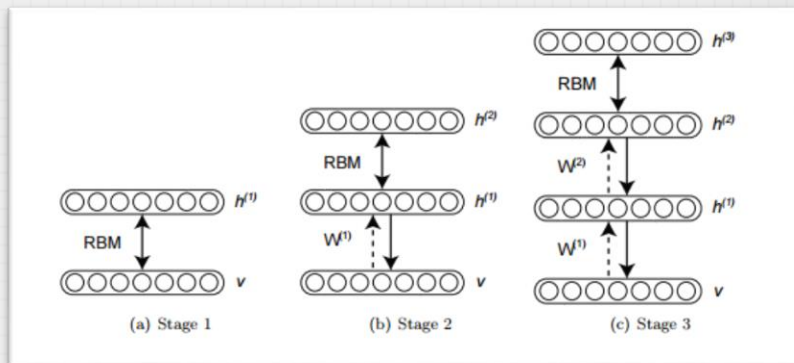


图片来自：Cortes, C.; Vapnik, V. (1995). "Support-vector networks". Machine Learning 20 (3): 273–297



三起： 2006之深度学习元年

- 无监督学习
- 分层预训练
- 新的网络结构
 - ▣ DBN、DBM、DAE
- 得名“深度”学习
- 优化方法的突破是第三次NN研究浪潮兴起的钥匙



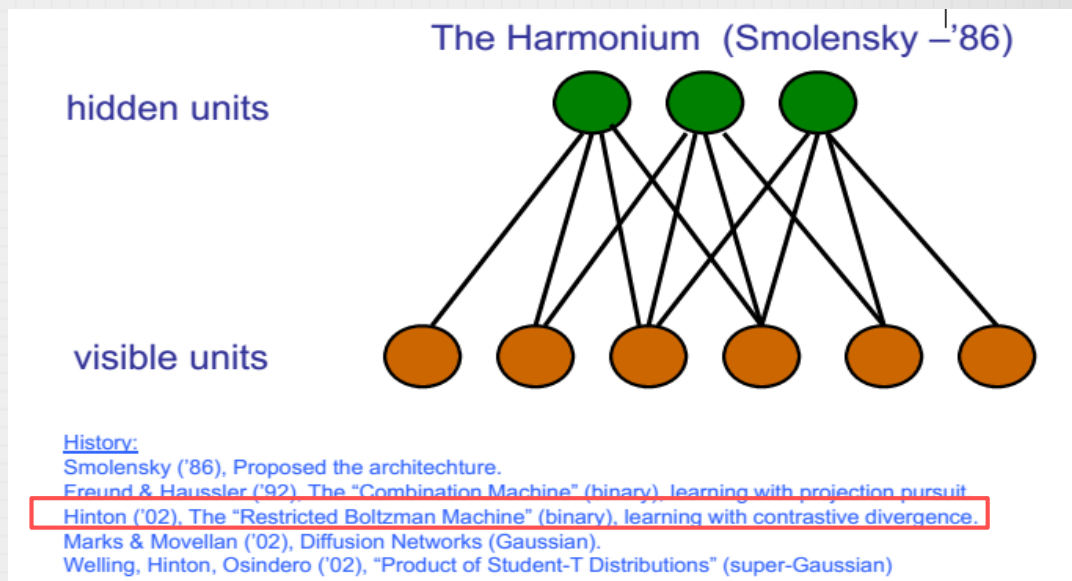
$$P(\mathbf{v}, \mathbf{h}^1, \mathbf{h}^2, \dots, \mathbf{h}^l) = P(\mathbf{v} | \mathbf{h}^1) P(\mathbf{h}^1 | \mathbf{h}^2) \dots P(\mathbf{h}^{l-2} | \mathbf{h}^{l-1}) P(\mathbf{h}^{l-1}, \mathbf{h}^l)$$

Hinton, G. E., Osindero, S. and Teh, Y., A fast learning algorithm for *deep belief nets*. Neural Computation 18:1527-1554, 2006
Hinton, G. E. and Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks. Science, Vol.313. no. 5786, pp. 504 - 507, 28 July 2006
Yoshua Bengio, Pascal Lamblin, Dan Popovici and Hugo Larochelle, Greedy *Layer-Wise Training of Deep Networks*, NIPS2006



三起：二十年磨一剑

- 1986年RBM架构提出
- 2002年Hinton找到快速学习算法
- 2006年DBN提出

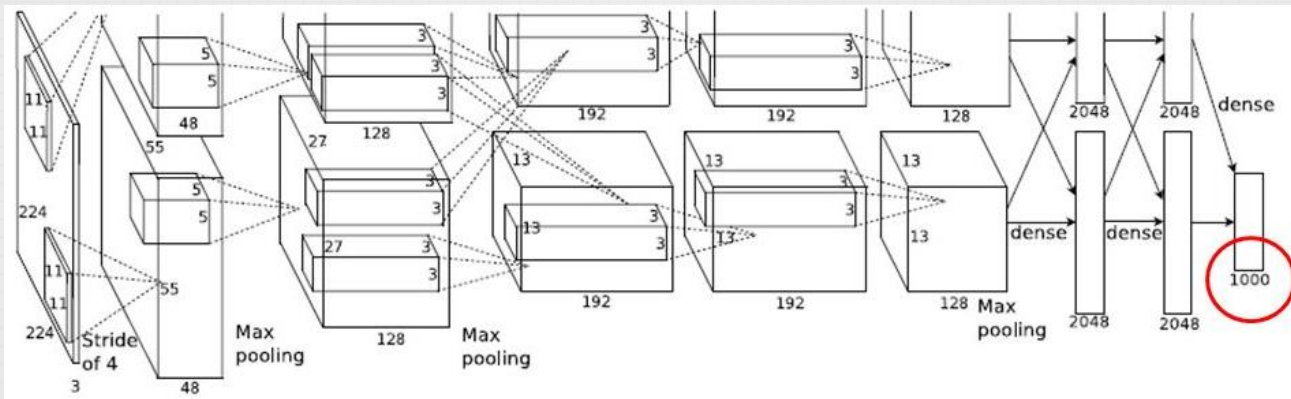


Credited to Prof. Eric Xing



三起：历史转折2012年ImageNet冠军AlexNet

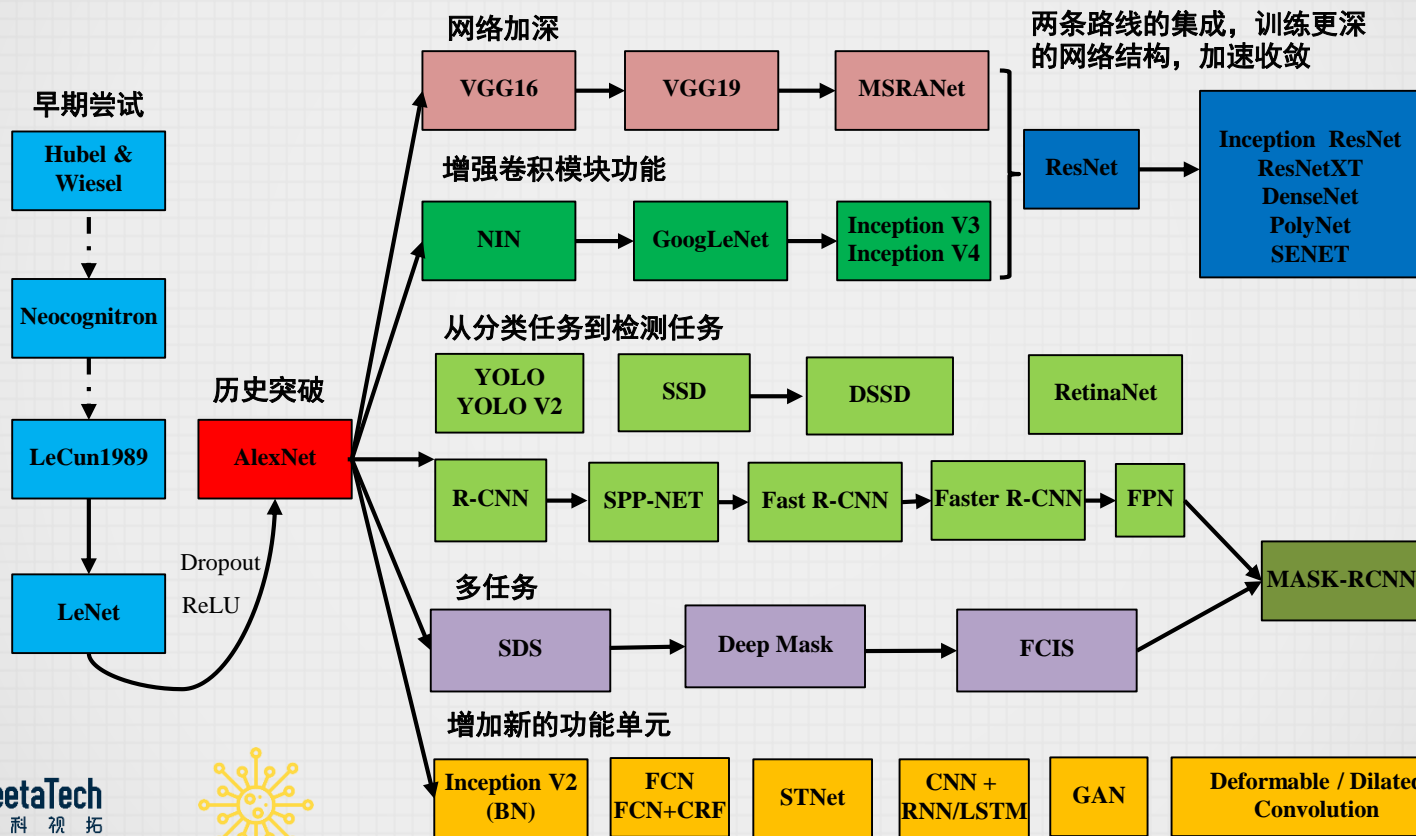
- 非线性激活函数：ReLU，克服了梯度消失
- 防止过拟合：Dropout, 数据增广
- 大数据训练：百万级ImageNet图像数据
- 其他：分Group实现双GPU并行，LRN归一化层
- 650K神经元，60M参数



Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.



三起：卷积神经网络结构演化



目录 CONTENTS

1

深度学习引导

2

深度学习简史

3

总结与发展



深度武功，四“大”不破

➤ 大数据

- ImageNet：百万级数据
- 中科视拓人脸识别模型训练：千万人、数以亿计的数据

➤ 大模型

- 从10层到1,001层
- Inception V3,V4等更复杂的结构

➤ 大算力

- GPU集群
- NVLink, InfiniBand

➤ 大社区

- 智力众包与协同创新



深度学习的未来发展

- 监督学习的红利是否消失？
 - ImageNet2016、2017竞赛再无方法上的惊喜
- 结构可学习
 - 设计网络结构是另外一种意义的Hand-crafted特征
 - 从人脑的结构中得到启发，例如跨层的连接、反馈机制
- 脏乱差数据学习
 - 摆脱对大而精确标注数据的依赖，利用标签有噪声或者缺失的数据
- 产生式模型的价值挖掘
 - GAN如何真正的在工业级应用中体现价值



Thank you!

