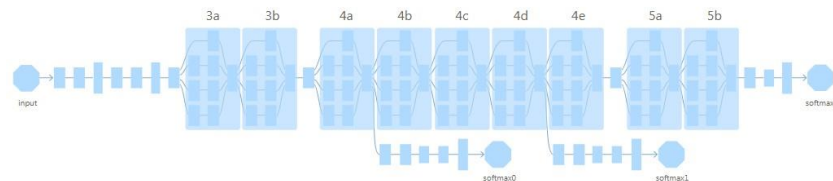




浙江大学城市学院
ZHEJIANG UNIVERSITY CITY COLLEGE



深度学习应用开发

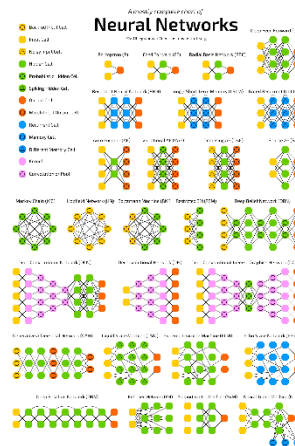
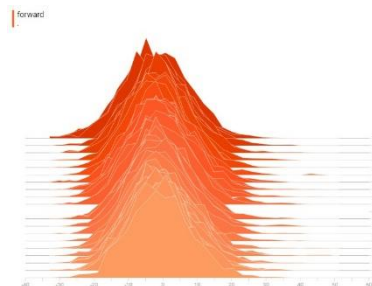
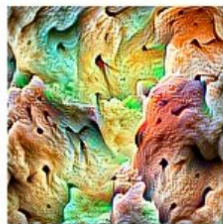
基于TensorFlow的实践

吴明晖 李卓蓉 金苍宏

浙江大学城市学院

计算机与计算科学学院

Dept. of Computer Science
Zhejiang University City College





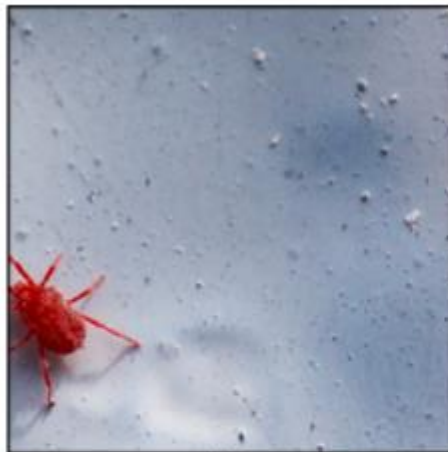
卷积神经网络 Convolutional Neural Network (CNN)



图像分类



浙江大學城市學院
ZHEJIANG UNIVERSITY CITY COLLEGE



mite



container ship



motor scooter



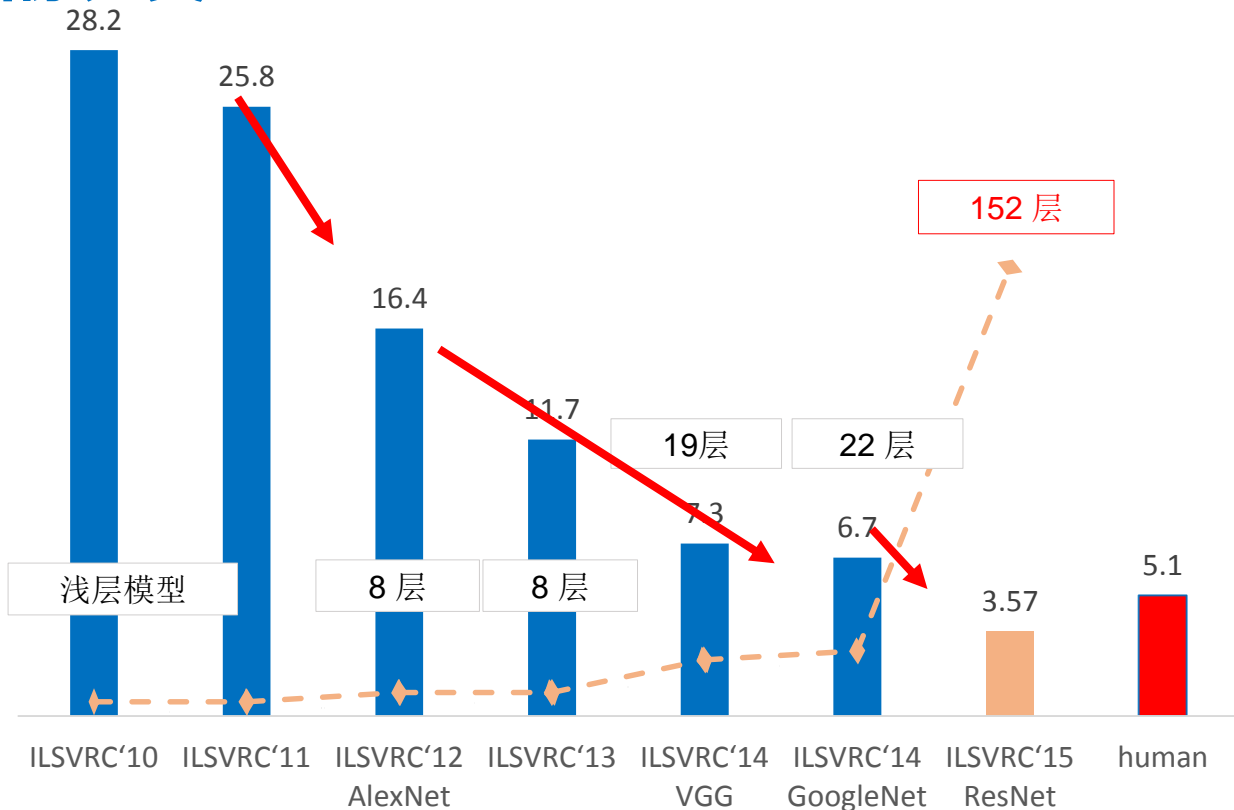
leopard

			
mite	container ship	motor scooter	leopard
black widow	lifeboat	go-kart	jaguar
cockroach	amphibian	moped	cheetah
tick	fireboat	bumper car	snow leopard
starfish	drilling platform	golfcart	Egyptian cat

Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton,
ImageNet Classification with Deep Convolutional Neural Networks, NIPS, 2012.



图像分类



ImageNet图像分类错误率 (top-5)



图像分类



浙江大學城市學院
ZHEJIANG UNIVERSITY CITY COLLEGE



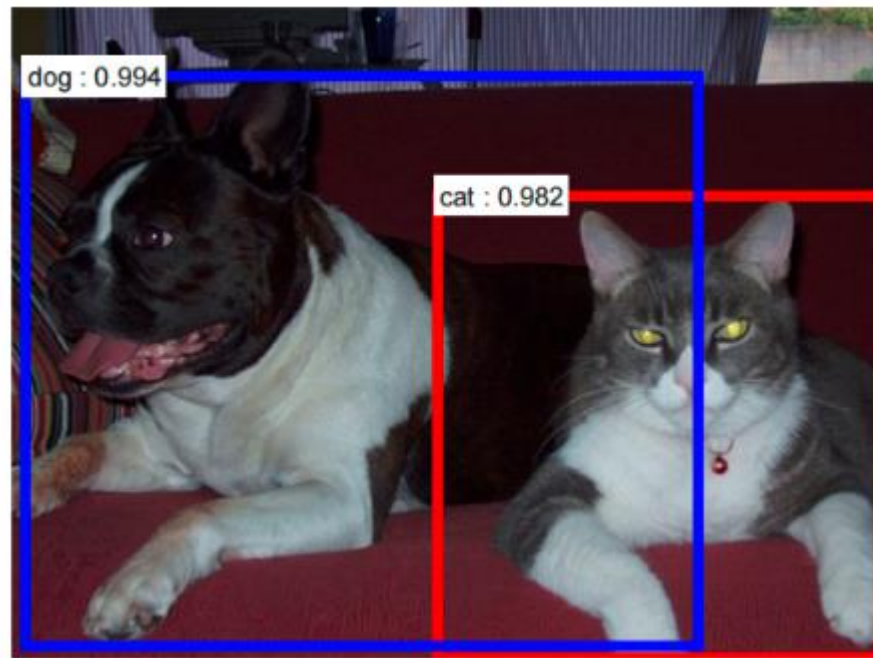
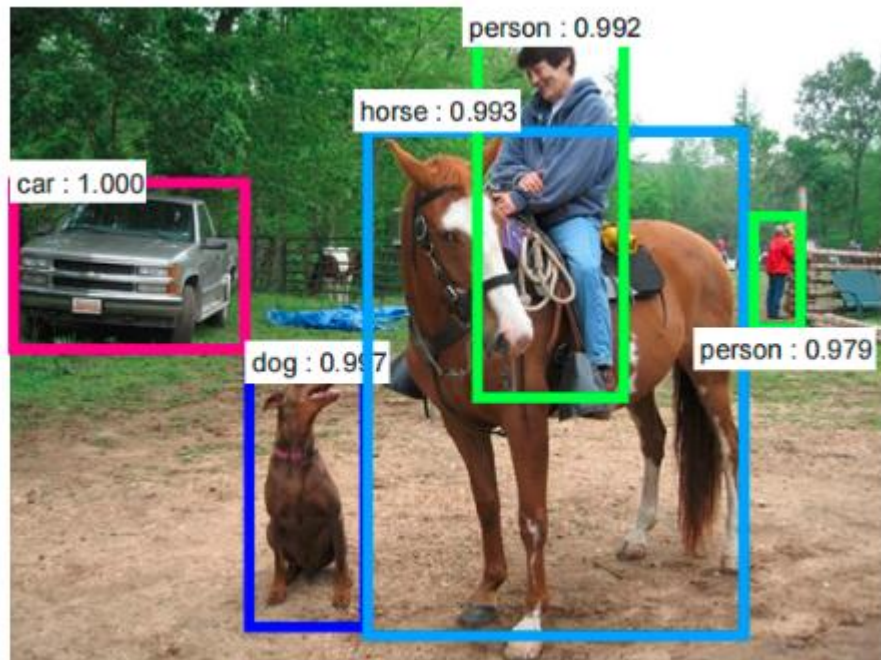
CIFAR-10数据集中分类难度较高的样例



物体检测



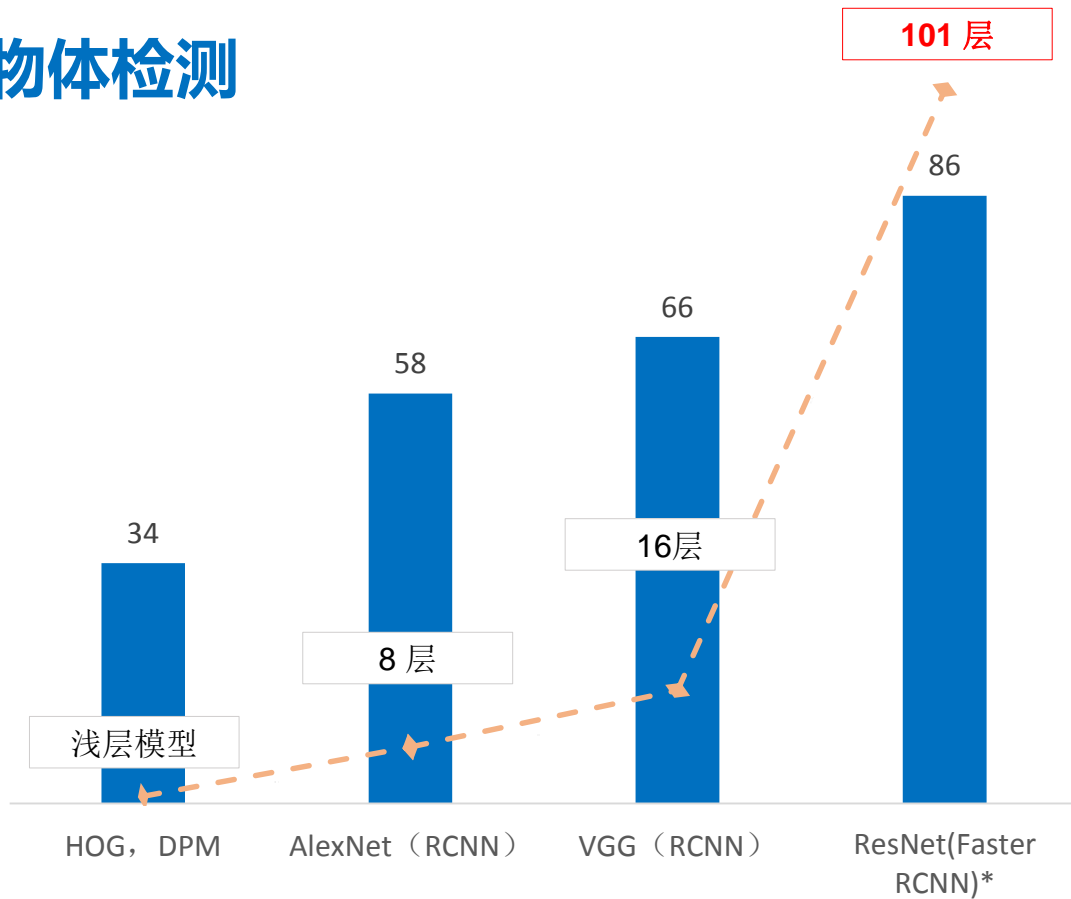
浙江大學城市學院
ZHEJIANG UNIVERSITY CITY COLLEGE



Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun,
Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.



物体检测



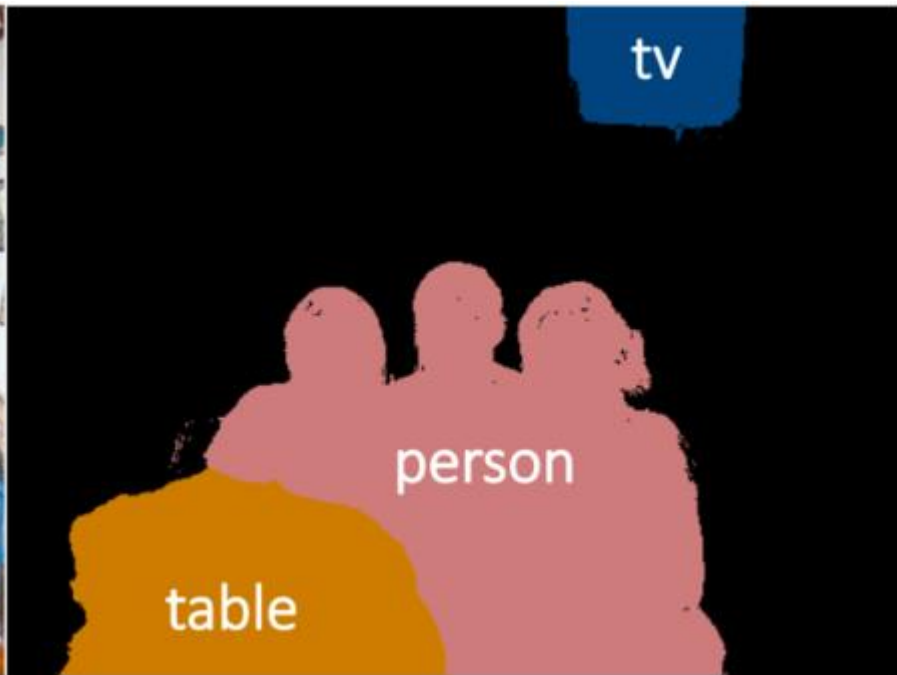
PASCAL VOC 2007 物体检测准确度



实例分割



浙江大學城市學院
ZHEJIANG UNIVERSITY CITY COLLEGE



Jifeng Dai, Kaiming He, Jian Sun,
BoxSup: Exploiting Bounding Boxes to Supervise Convolutional Networks for Semantic Segmentation.



看图说话



浙江大學城市學院
ZHEJIANG UNIVERSITY CITY COLLEGE



man in black shirt is playing guitar.



construction worker in orange safety vest is working on road.



two young girls are playing with lego toy.



boy is doing backflip on wakeboard.

Andrej Karpathy, Li Fei-Fei,
Deep Visual-Semantic Alignments for Generating Image Description, CVPR.



看图问答



What kind of store is this?	bakery	art supplies
	bakery	grocery
	pastry	grocery
Is the display case as full as it could be?	no	no
	no	yes
	no	yes



How many bikes are there?	2	3
	2	4
	2	12
What number is the bus?	48	4
	48	46
	48	number 6



内容安排



浙江大学城市学院
ZHEJIANG UNIVERSITY CITY COLLEGE

卷积神经网络成了机器视觉任务的一个极为重要的工具！



卷积神经网络的结构



Tensorflow对卷积神经网络的支持



案例：CIFAR-10图像识别



全连接神经网络的局限性



全连接网络的局限性

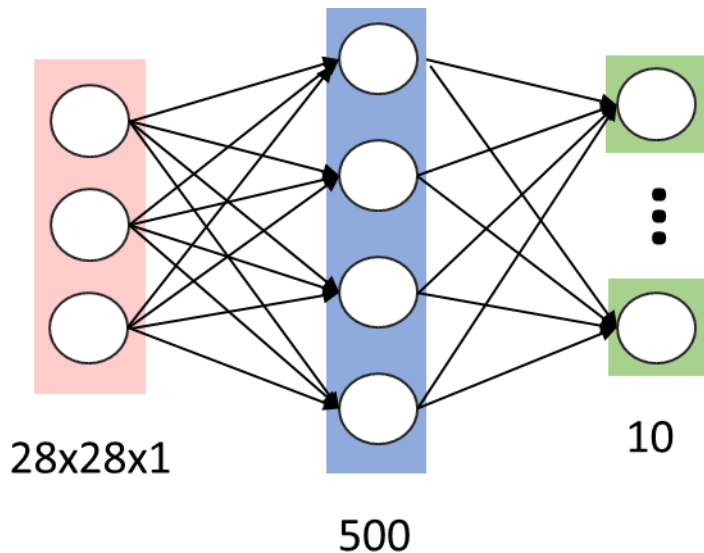


对于MNIST 手写数字识别，假如第一个隐层的节点数为500，那么一个全连接层的参数个数为：

➤ $28 \times 28 \times 1 \times 500 + 500 \approx 40$ 万；

当图片更大时，比如CIFAR数据集中，图片大小为 $32 \times 32 \times 3$ ，此时全连接层的参数有：

➤ $32 \times 32 \times 3 \times 500 + 500 \approx 150$ 万





全连接网络的局限性

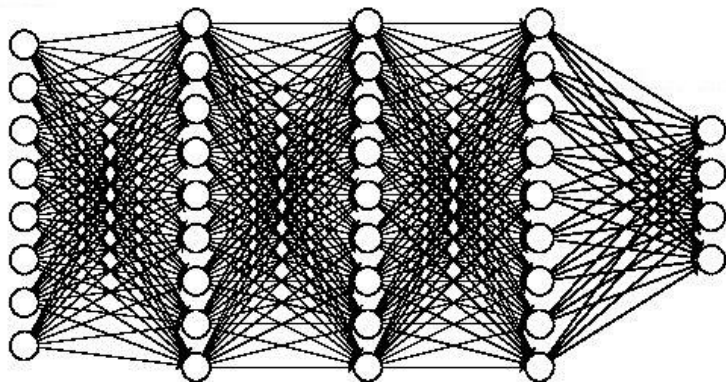


当图片分辨率进一步提高时，当隐层数量增加时，
例如：600 x 600 图像，各隐层节点数分别为
300,200和100，则参数个数为：

$$600 \times 600 \times 300 + 300 \times 200 + 200 \times 100 \\ \approx 1.08\text{亿}$$

参数增多会导致：

- 计算速度减慢
- 过拟合



需要更合理的结构来有效减少参数个数！