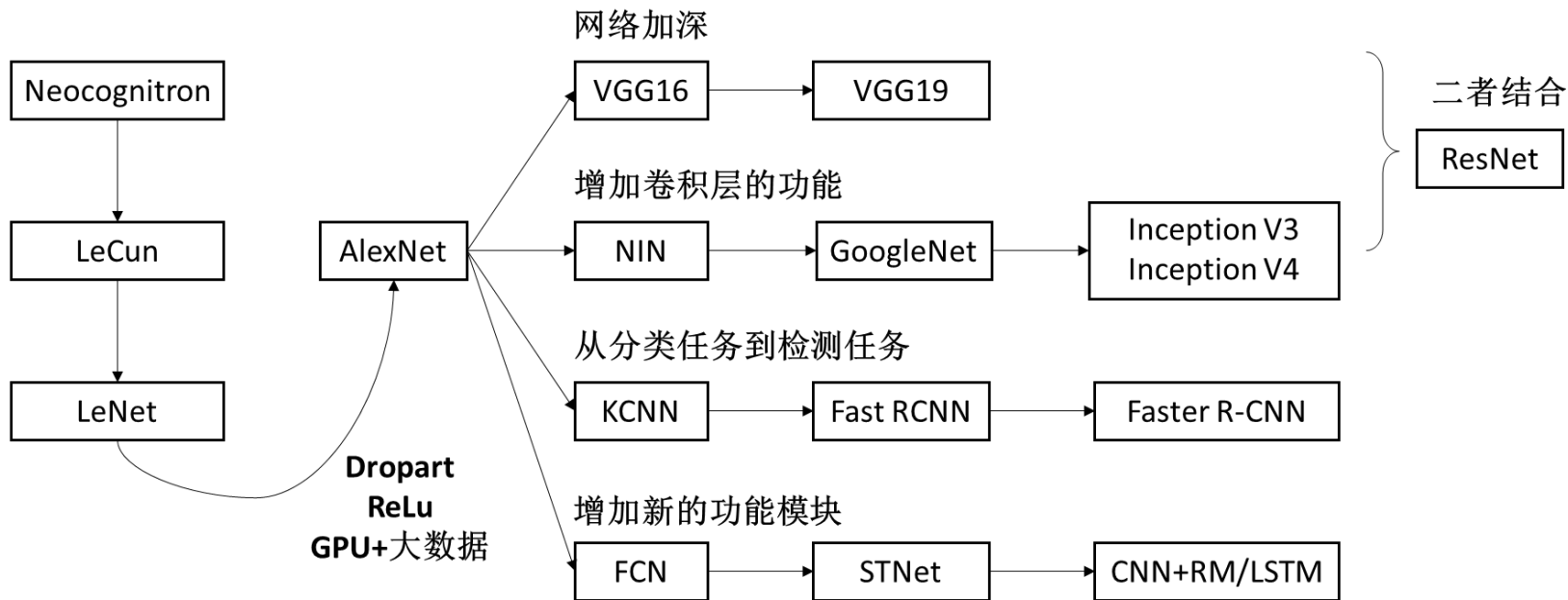




经典CNN





经典CNN——AlexNet



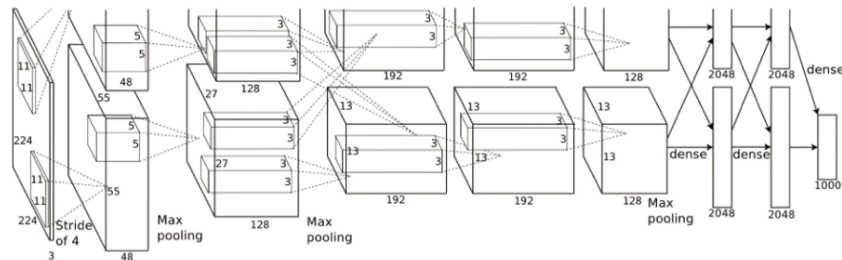
浙江大學城市學院
ZHEJIANG UNIVERSITY CITY COLLEGE

- 2012年ImageNet LSVRC（大规模视觉识别挑战赛）**冠军**

- TOP-5错误率15.3%，远低于第二名的错误率（26.2%）；TOP-1错误率37.5%

- 主要贡献：

- 防止过拟合：数据增强（data augmentation），Dropout
 - GPU实现：将网络分布在两个GPU上，且GPU之间在某些层能够互相通信。
 - 非线性激活：ReLU
 - 大数据训练：120万ImageNet图像数据集





数据增强



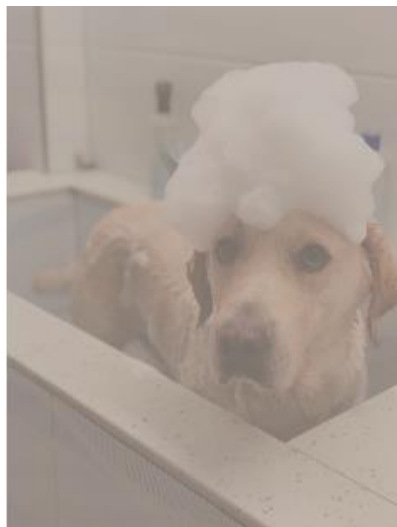
浙江大學城市學院
ZHEJIANG UNIVERSITY CITY COLLEGE



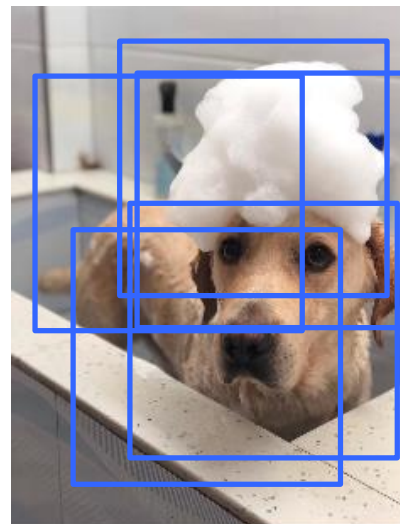
(a) 原图



(b) 水平翻转



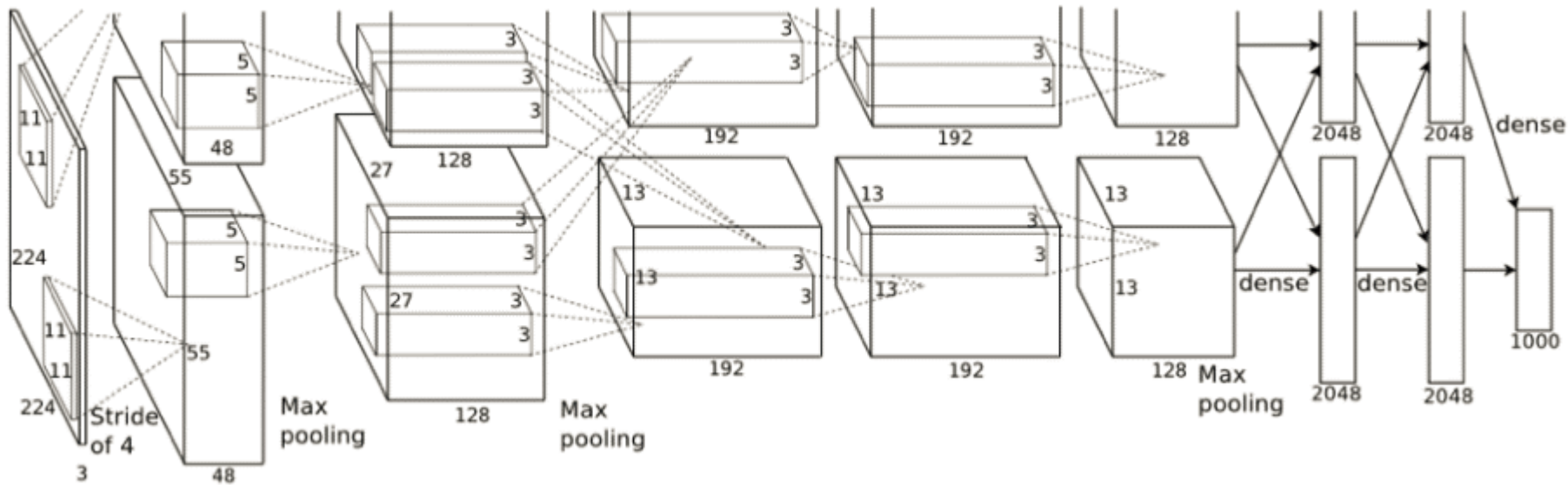
(c) 改变对比度



(d) 随机裁剪

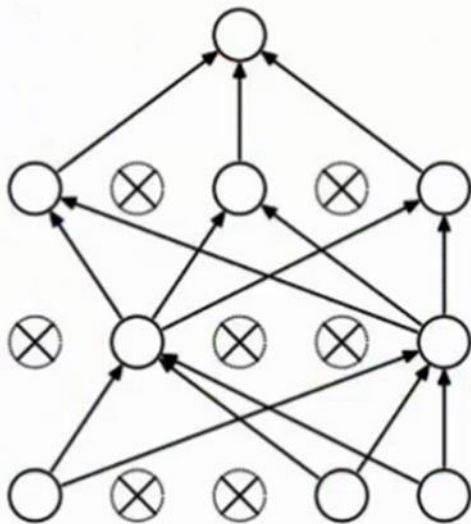
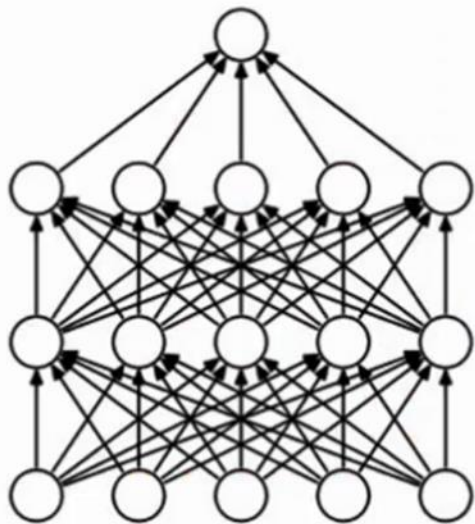


经典CNN——AlexNet





dropout



- 用于全连接层
- 每次迭代以某概率将神经元输出置零，不参与前向和反向传播
- 产生不同的网络结构，进行组合，大大减少了过拟合
- 缺点：训练时间增加



经典CNN——VGGNet

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

①

②

③

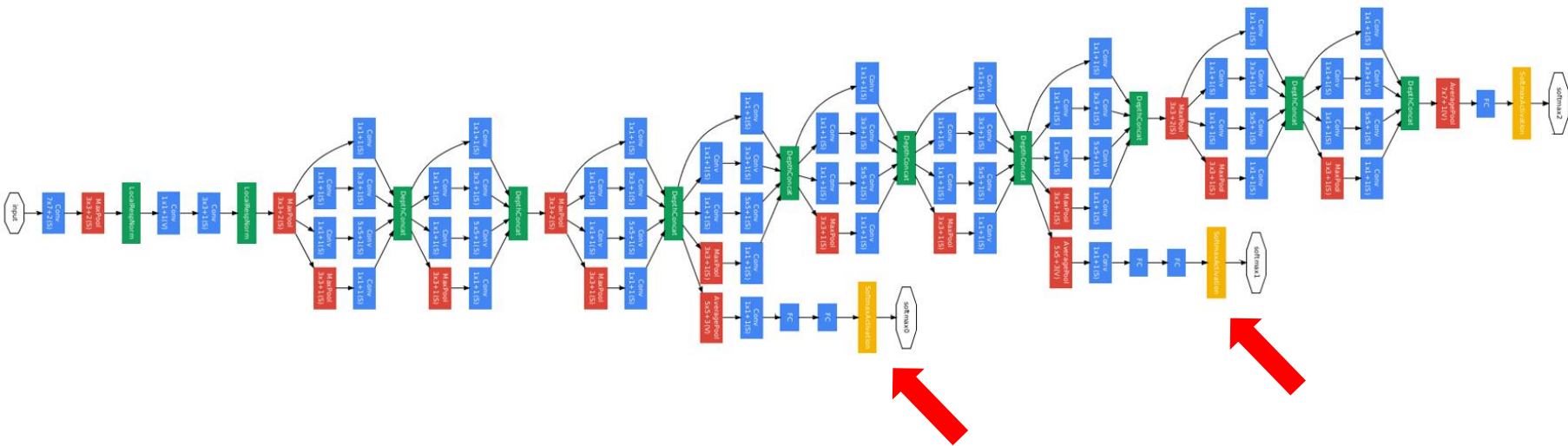
④

⑤

- 可以看成加深版本的AlexNet
- 5个卷积组
- Conv X-Y: 卷积核尺寸X, 深度Y
- 11层~19层
- 卷积核大小: 3x3
- 卷积核深度: 大部分都采用了逐层递增的方式
- VGG-16和VGG-19



经典CNN——GoogleNet



- 深度
- 宽度

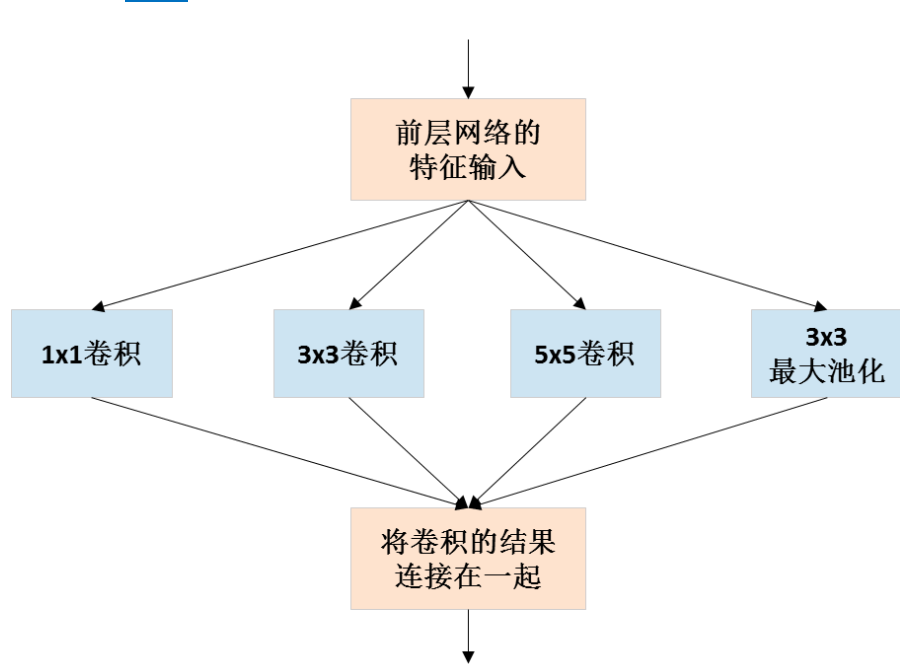
Szegedy C, Liu W, Jia Y, et al.
Going deeper with convolutions.



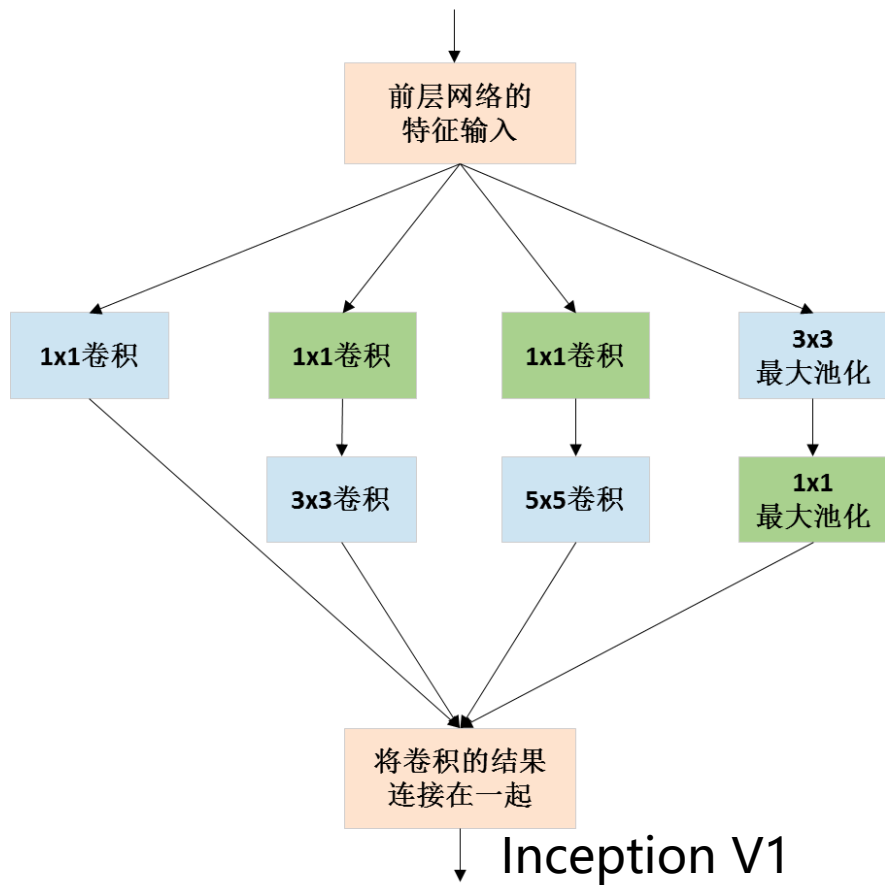
经典CNN——GoogleNet



浙江城市学院
ZHEJIANG UNIVERSITY CITY COLLEGE

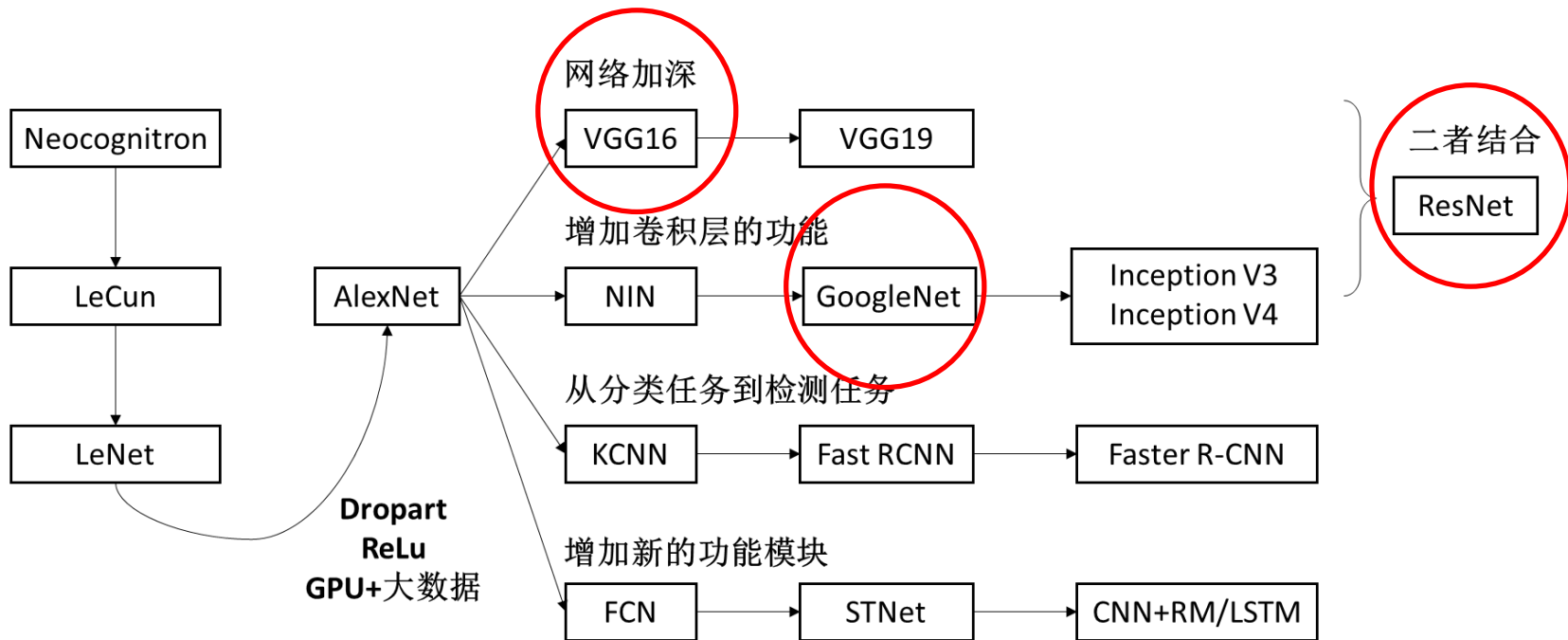


Inception V0



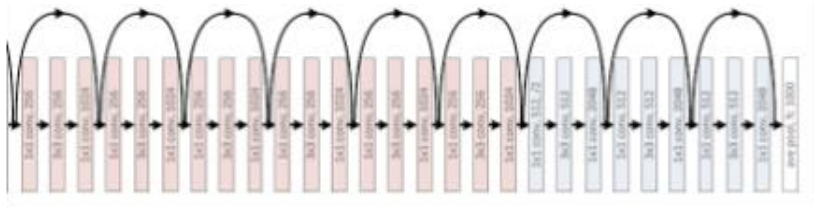


经典CNN——ResNet



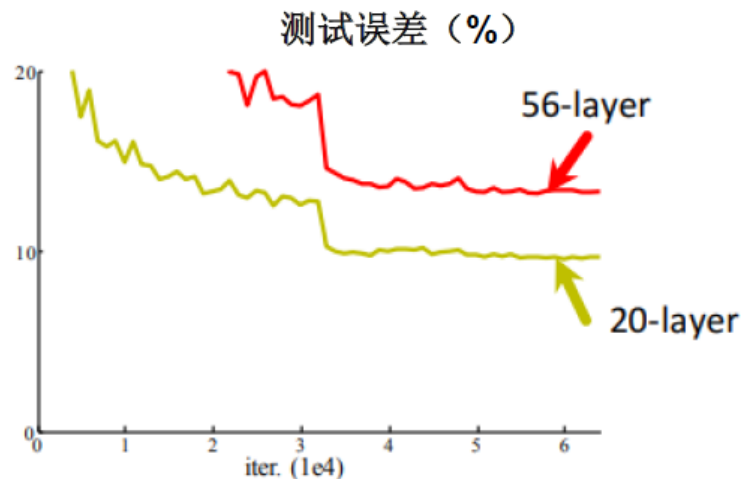
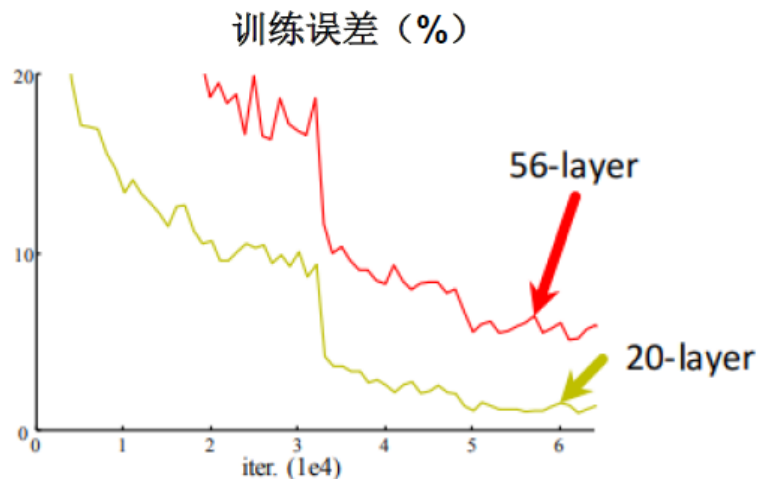


经典CNN——ResNet





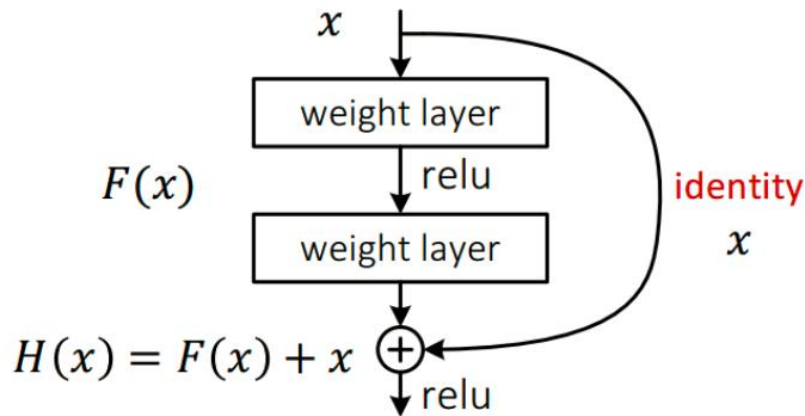
经典CNN——ResNet



- 56层网络比20层网络有更高的训练和测试误差
- 非过拟合问题
- 梯度消失导致无法对前面网络层的权重进行有效调整



经典CNN——ResNet



捷径连接 (shortcut connections)

- 学习一个残差映射： $x \rightarrow F(x) = H(x) - x$
- 而不是复杂映射： $x \rightarrow F(x) = H(x)$
- 学习 $F(x) = 0$ 相较于学习 $H(x) = x$ 更容易