



# 残差神经网络

王维率

# Resnet分类网络

- 当前应用最为广泛的CNN特征提取网络
- 于2015年提出
- He-Kaiming, Ren-Shaoqing, Sun-Jian

# 起源

“退化问题就暴露了：随着神经网络深度的增加，精确度开始饱和（这是不足为奇的），然后会迅速地变差。”

- 一般印象：深度学习愈是深（复杂，参数多）愈是有着更强的表达能力
- CNN分类网络：7层（Alexnet）→ 16层、19层（VGG）→ 22层（Googlenet）
- 问题：深度CNN网络达到一定深度后再一味地增加层数并不能带来进一步地分类性能提高
- 发现：过深的网络仍然还会使分类准确度下降

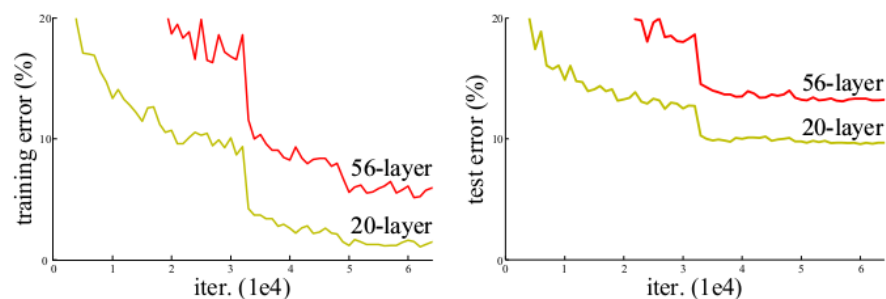


Figure 1. Training error (left) and test error (right) on CIFAR-10 with 20-layer and 56-layer “plain” networks. The deeper network has higher training error, and thus test error. Similar phenomena on ImageNet is presented in Fig. 4.

# 残差学习

- 一般的CNN网络:  $X \rightarrow H(X)$
- 残差学习:  $H(X) \rightarrow X$   
 $X \rightarrow (H(X) - X) + X$
- 使用多个有参网络层来学习输入、输出之间的残差
- $X$ : 直接的identity mapping
- $H(X) - X$ : 有参网络层要学习的输入输出间残差

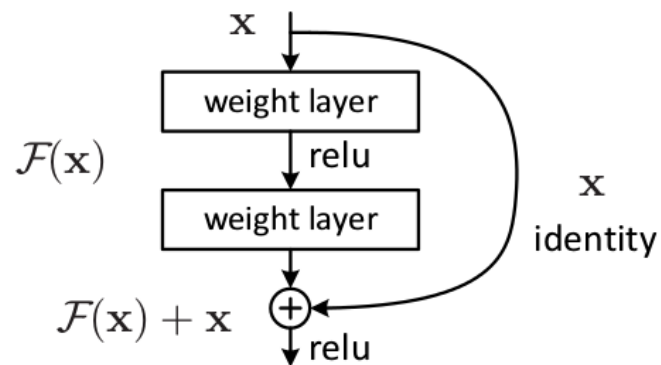


Figure 2. Residual learning: a building block.

# Identity mapping

- 在输入、输出之间建立了一条直接的关联通道  
从而使得强大的有参层集中精力学习输入、输出之间的残差
- $F(X, W_i)$ 来表示残差映射  
输出即为:  $Y = F(X, W_i) + X$
- 输入、输出通道数相同时: 直接相加
- 输入、输出通道数不同时:
  - 简单地将X相对Y缺失的通道直接补零从而使其能够相对齐的方式
  - 通过使用1x1的conv来表示 $W_s$ 映射从而使得最终输入与输出的通道达到一致的方式

# 残差网络结构与朴素网络结构

- 共使用了三种网络进行实验：
  - VGG19网络
  - VGG朴素网络
  - 与上述朴素网络相对应的Resnet网络
- 对于其中的输入、输出通道数目不同的情况使用两种可能的选择：
  - shortcut直接使用identity mapping，不足的通道通同补零来对齐
  - 使用1x1的Conv来表示 $W_s$ 映射，从而使得输入、输出通道数目相同

# 残差网络结构与朴素网络结构

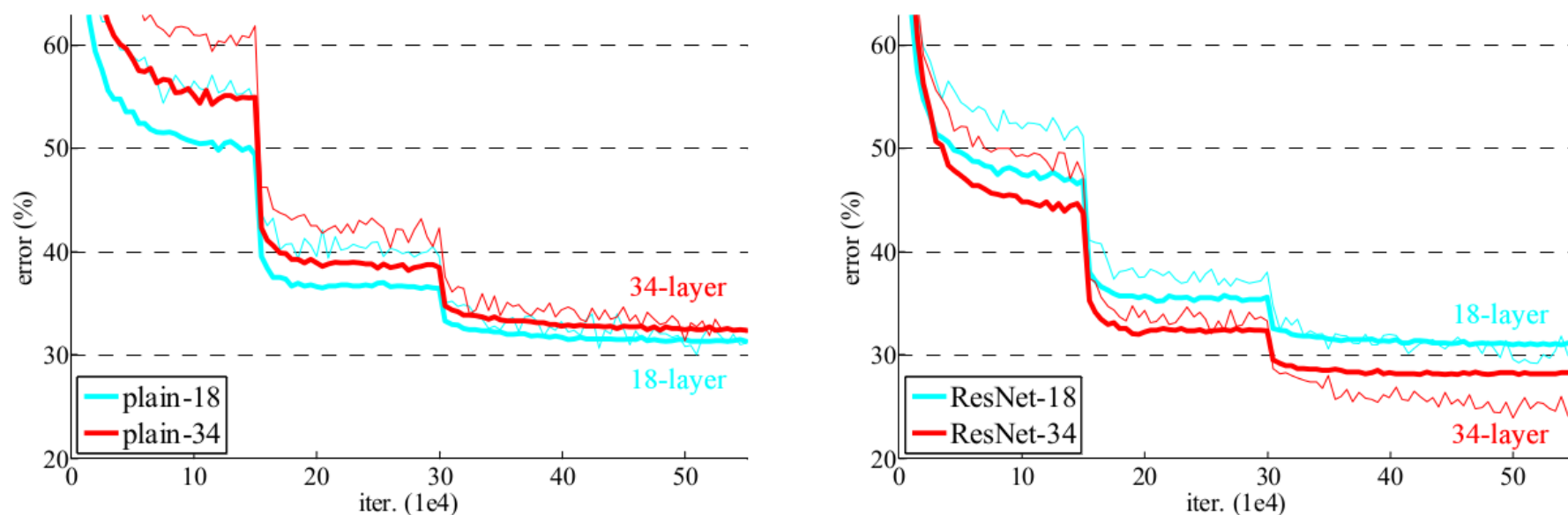


Figure 4. Training on **ImageNet**. Thin curves denote training error, and bold curves denote validation error of the center crops. Left: plain networks of 18 and 34 layers. Right: ResNets of 18 and 34 layers. In this plot, the residual networks have no extra parameter compared to their plain counterparts.

# 残差网络结构与朴素网络结构

	plain	ResNet
18 layers	27.94	27.88
34 layers	28.54	<b>25.03</b>

Table 2. Top-1 error (% , 10-crop testing) on ImageNet validation. Here the ResNets have no extra parameter compared to their plain counterparts. Fig. 4 shows the training procedures.



# CIFAR10上更深的Resnet网络

- 疯狂的实验：1202层
- 实验结果：效果终于下降
- 尝试分析：层数过多，导致模型过于复杂，  
而CIFAR-10较小的数据集造成了它的过拟合

method			error (%)
Maxout [9]			9.38
NIN [25]			8.81
DSN [24]			8.22
	# layers	# params	
FitNet [34]	19	2.5M	8.39
Highway [41, 42]	19	2.3M	7.54 (7.72±0.16)
Highway [41, 42]	32	1.25M	8.80
ResNet	20	0.27M	8.75
ResNet	32	0.46M	7.51
ResNet	44	0.66M	7.17
ResNet	56	0.85M	6.97
ResNet	110	1.7M	<b>6.43</b> (6.61±0.16)
ResNet	1202	19.4M	7.93

Table 6. Classification error on the **CIFAR-10** test set. All methods are with data augmentation. For ResNet-110, we run it 5 times and show “best (mean±std)” as in [42].



谢谢大家！