ResNet (Residual Networks) is a deep learning architecture introduced by Kaiming He et al. in 2015 through their research paper "Deep Residual Learning for Image Recognition." It was designed to tackle the vanishing gradient problem in very deep neural networks. ResNet played a crucial role in enabling deep networks with hundreds or even thousands of layers while maintaining training stability and accuracy.

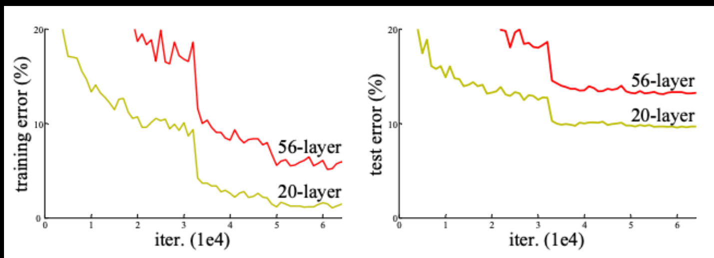
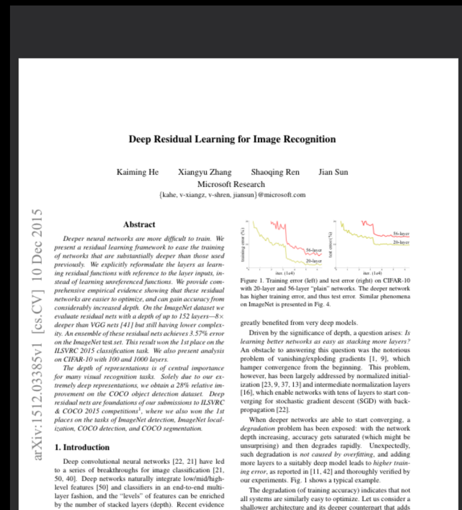


Figure 1. Training error (left) and test error (right) on CIFAR-10 with 20-layer and 56-layer "plain" networks. The deeper network has higher training error, and thus test error. Similar phenomena on ImageNet is presented in Fig. 4.


Deep Residual Learning for Image Recognition

Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun
Microsoft Research

arXiv:1512.03385v1 [cs.CV] 10 Dec 2015

Abstract

1. Introduction

Agenda

1. Network Architecture
2. Residual Learning
3. Resnet Variants
4. Solving Vanishing Gradient Problem

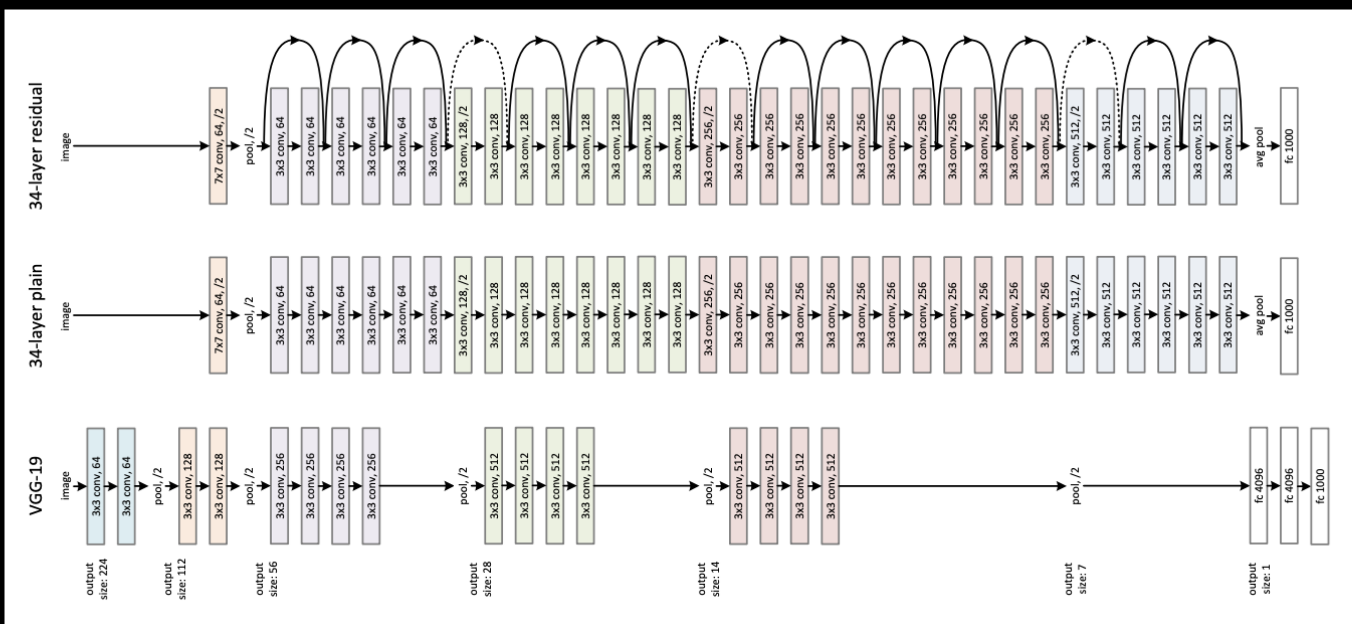


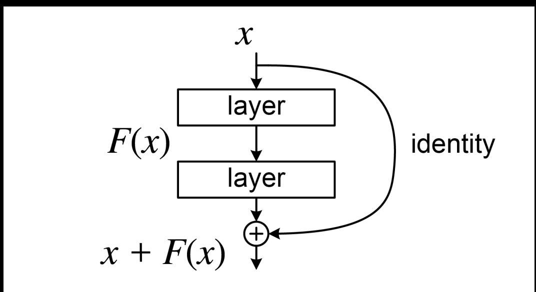
Figure 3. Example network architectures for ImageNet. **Left**: the VGG-19 model [41] (19.6 billion FLOPs) as a reference. **Middle**: a plain network with 34 parameter layers (3.6 billion FLOPs). **Right**: a residual network with 34 parameter layers (3.6 billion FLOPs). The dotted shortcuts increase dimensions. **Table 1** shows more details and other variants.

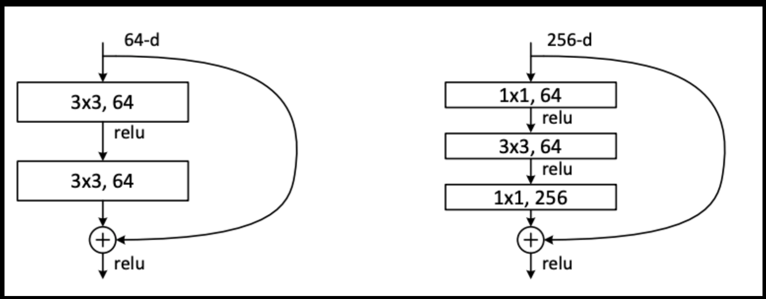
1) Feature Reuse

3) Degradation

2) Solving VGD

Residual Blocks





$x + f(x)$

1) Basic Residual Block → Resnet 34

2) Residual Bottle neck Block → Resnet 50, 101, 152

Structure :- 2 or 3 Conv Layers

BRB

Batch Norm

Relu Act

Skip Connections

RBB Structure :- 1×1 Conv

3×3 Conv

1×1 Conv

computationally efficient

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix}3\times3, 64\\3\times3, 64\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 64\\3\times3, 64\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 64\\3\times3, 64\\1\times1, 256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 64\\3\times3, 64\\1\times1, 256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 64\\3\times3, 64\\1\times1, 256\end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix}3\times3, 128\\3\times3, 128\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 128\\3\times3, 128\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, 128\\3\times3, 128\\1\times1, 512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, 128\\3\times3, 128\\1\times1, 512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, 128\\3\times3, 128\\1\times1, 512\end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix}3\times3, 256\\3\times3, 256\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 256\\3\times3, 256\end{bmatrix}\times6$	$\begin{bmatrix}1\times1, 256\\3\times3, 256\\1\times1, 1024\end{bmatrix}\times6$	$\begin{bmatrix}1\times1, 256\\3\times3, 256\\1\times1, 1024\end{bmatrix}\times23$	$\begin{bmatrix}1\times1, 256\\3\times3, 256\\1\times1, 1024\end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix}3\times3, 512\\3\times3, 512\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 512\\3\times3, 512\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 512\\3\times3, 512\\1\times1, 2048\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 512\\3\times3, 512\\1\times1, 2048\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 512\\3\times3, 512\\1\times1, 2048\end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Input = (224,224,3)

General Structure

1) Conv + MP (7x7)
2) Multiple Residual Blocks
3) GAP
4) FC
5) Softmax

Solving VGD

1) Skip Connections

2) Identity Mapping

ResNet 50

Stages	Layers
Conv 1	7x7 Conv, S2 + MP
Stage 1	3 x BottleNeck (1x1, 3x3, 1x1)
Stage 2	4 x BottleNeck
Stage 3	6 x BottleNeck
Stage 4	3 x BottleNeck
GAP	
FC with Softmax	