

ResNet (CVPR 2016)

Deep Residual Learning for Image Recognition

Kaiming He
etc.

① Problem Description:

Learning better with more layers?

Problem {
vanishing/exploding gradients
degradation problem

Exist Solution: Normalization

② Problem Solution:

Deep Residual Learning:

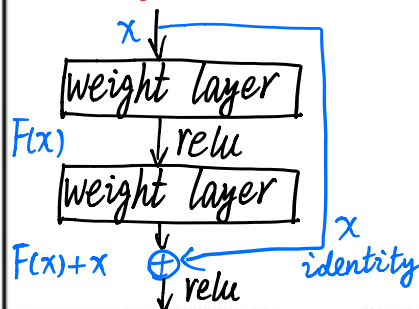
Let $H(x) = F(x) + x$;

$F(x)$ denotes "shortcut connections"

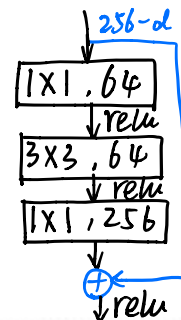
{
easy to optimize
easily enjoy accuracy gains from greatly
increased depth

③ Conceptual
Understanding

building block:



bottleneck:



building block for ResNet-34, bottleneck for ResNet-50/101/152.

(input and output)

same dimensions: identity shortcuts can be directly used.

dimensions {
increase { shortcut \rightarrow identity mapping, zero-padding \rightarrow increase dimension.
projection shortcuts \rightarrow match dimensions.

$[1 \times 1]$: reducing and then increasing (restoring) dimensions

$[3 \times 3]$: leaving the 3×3 layer a bottleneck with smaller input/output dimensions

A blank sheet of graph paper featuring a grid of dashed horizontal and vertical lines. The grid consists of 6 columns and 8 rows, creating a total of 48 rectangular cells. A solid black border frames the entire page, with a slightly thicker line at the top edge.