ResNet (CVPR 2016)

Deep Residual Learning for Image Recognition

Kaiming He etc.

①Problem Description:

Learning better with more layers?

Problem { vanishing/exploding gradients
degradation problem

Exist Solution: Normalization

②Problem Solution:

Deep Residual Learning:

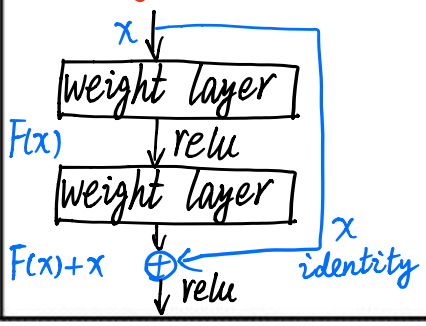
Let $H(x) = F(x) + x$;

F(x) denotes "shortcut connections"

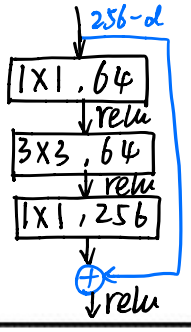
{ easy to optimize
easily enjoy accuracy gains from greatly increased depth

③Conceptual Understanding

building block:



bottleneck:



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

(input and ouput)

same dimensions: identity shortcuts can be directly used.

dimensions increase { Shotcut → identity mapping, zero-padding → increase dimension.
projection shotcuts → match dimensions.

1x1: reducing and then increasing (restoring) dimensions

3x3: leaving the 3×3 layer a bottleneck with smaller input/output dimensions