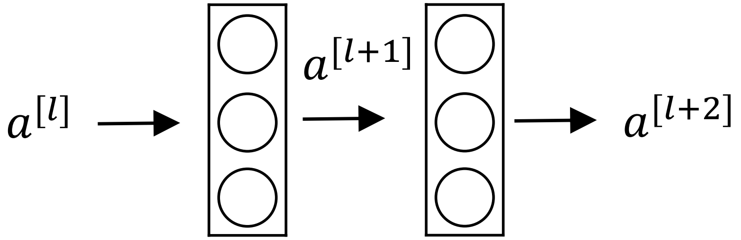
deeplearning.ai

Case Studies

Residual Networks (ResNets)

Residual block



Short cut (skip connection)

$a^{[l]} \to$ linear $\to$ Relu $\to a^{[l+1]} \to$ linear $\oplus \to$ Relu $\to a^{[l+2]}$

$\Rightarrow a^{[l+2]} = g(z^{[l+2]} + a^{[l]})$

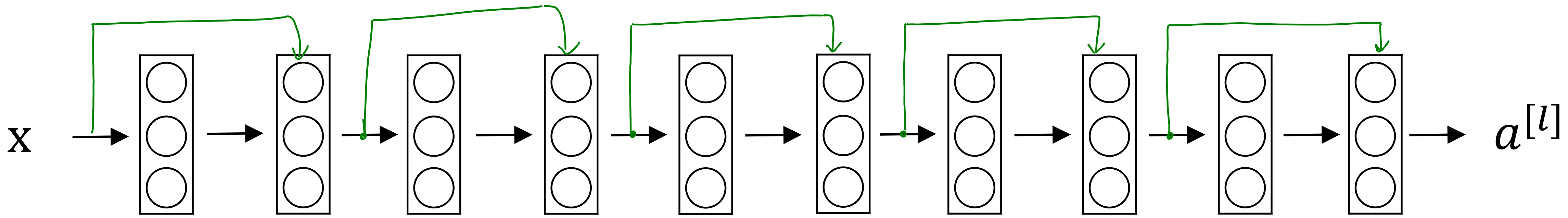
$z^{[l+1]} = W^{[l+1]} a^{[l]} + b^{[l+1]}$ (linear) $\quad a^{[l+1]} = g(z^{[l+1]})$ (Relu) $\quad z^{[l+2]} = W^{[l+2]} a^{[l+1]} + b^{[l+2]}$ (linear) $\quad a^{[l+2]} = g(z^{[l+2]})$ (Relu)

$\therefore$ to go from $a^{[l]}$ to $a^{[l+2]}$, we need to go through this "main path"

[He et al., 2015. Deep residual networks for image recognition]

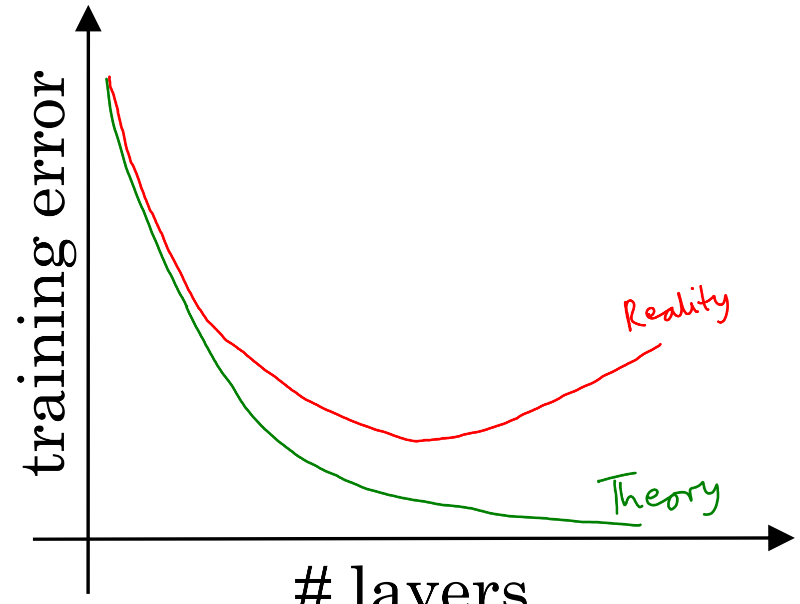
Andrew Ng

Residual Network

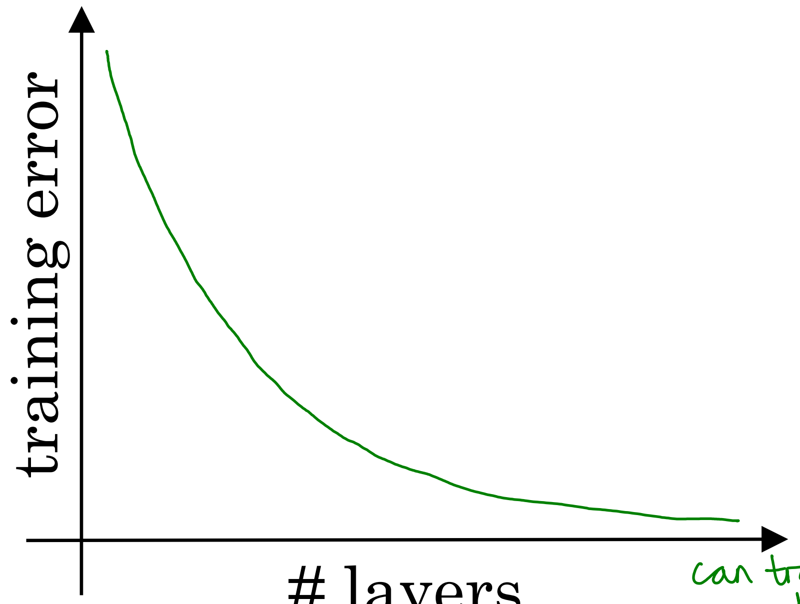


These are 5 Residual Blocks
stacked together to form a Residual N/W

Plain



ResNet



can train
very deep N/w
(≥ 150 layers)

[He et al., 2015. Deep residual networks for image recognition]

Andrew Ng