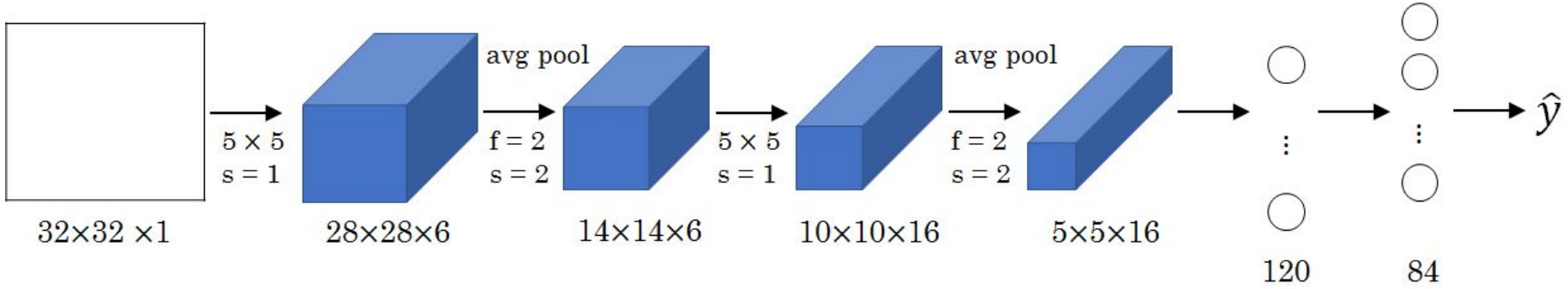
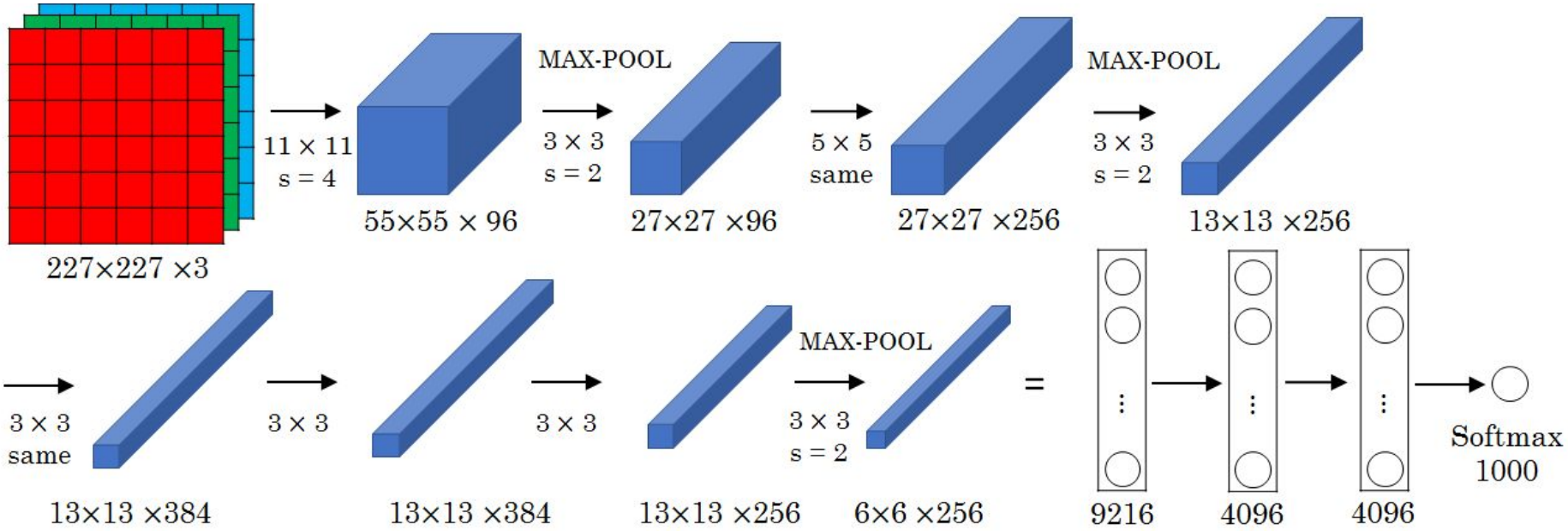
Classic Networks

LeNet - 5



[LeCun et al., 1998. Gradient-based learning applied to document recognition]

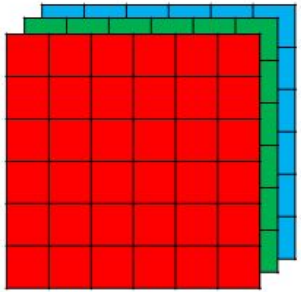
Alex Net



[Krizhevsky et al., 2012. ImageNet classification with deep convolutional neural networks]

VGG - 16

CONV = 3×3 filter, s = 1, same MAX-POOL = 2×2 , s = 2



224×224 ×3

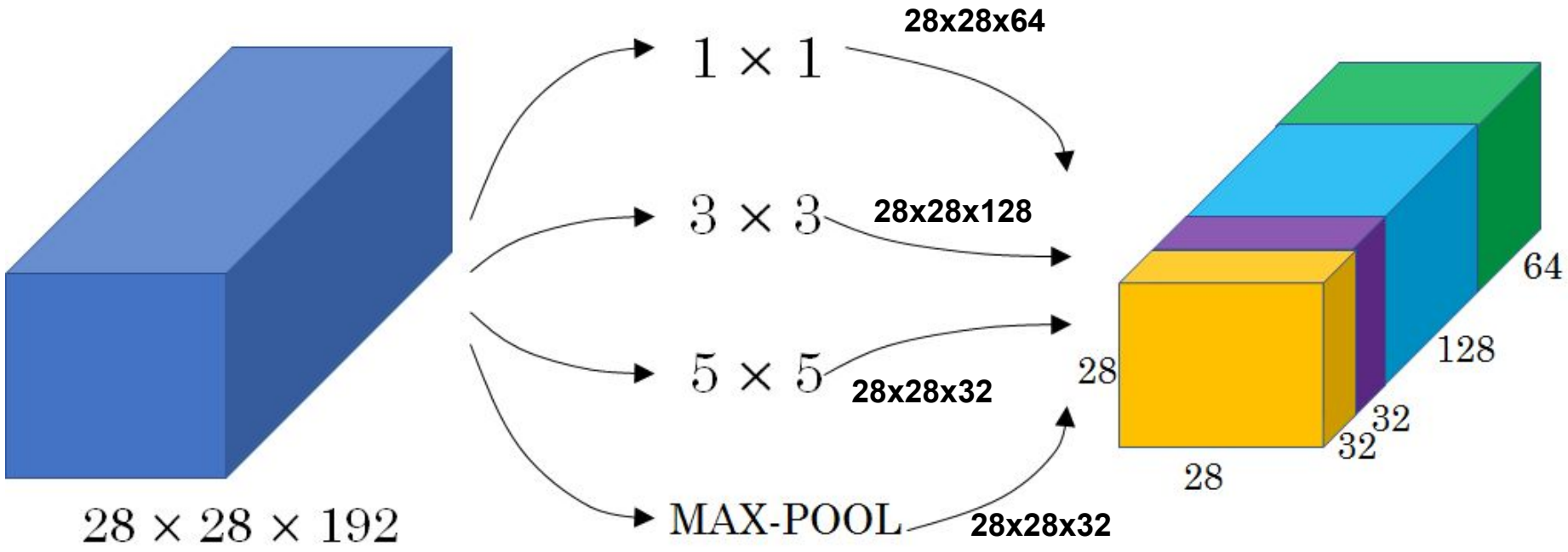
→ [CONV 64] ×2 → 224×224×64 → POOL → 112×112 ×64 → [CONV 128] ×2 → 112×112 ×128 → POOL → 56×56 ×128

→ [CONV 256] ×3 → 56×56 ×256 → POOL → 28×28 ×256 → [CONV 512] × 3 → 28×28 ×512 → POOL → 14×14×512

→ [CONV 512] ×3 → 14×14 ×512 → POOL → 7×7×512 → FC 4096 → FC 4096 → Softmax 1000

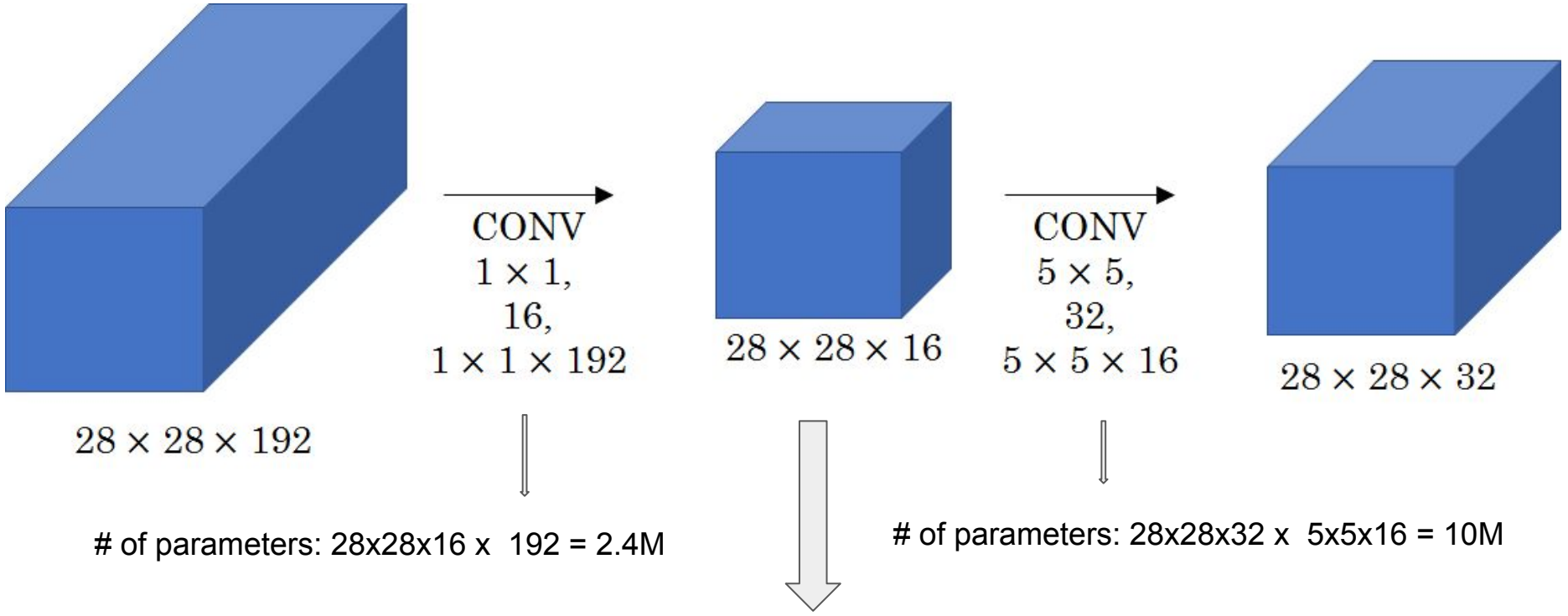
[Simonyan & Zisserman 2015. Very deep convolutional networks for large-scale image recognition]

Motivation for inception network

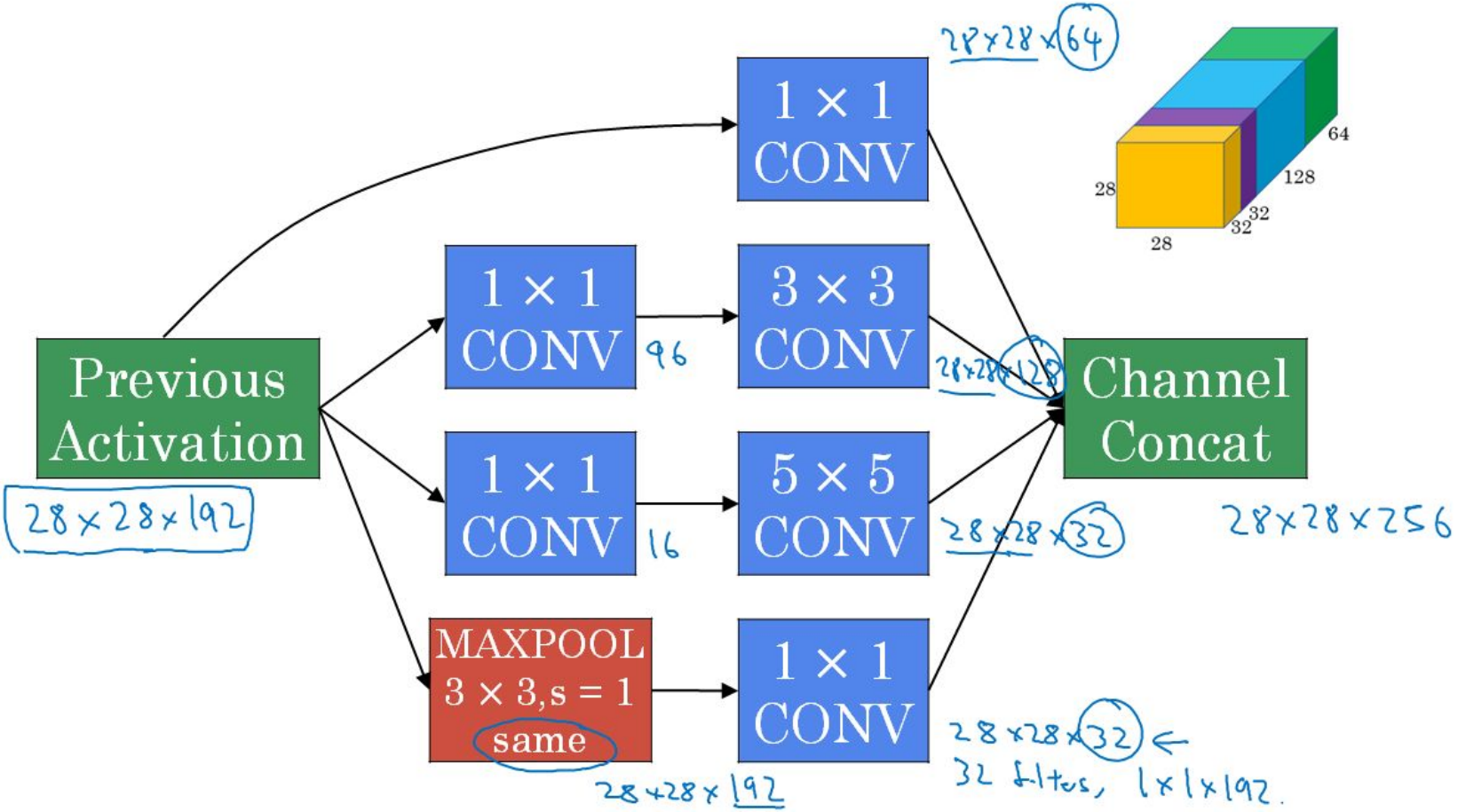


[Szegedy et al. 2014. Going deeper with convolutions]

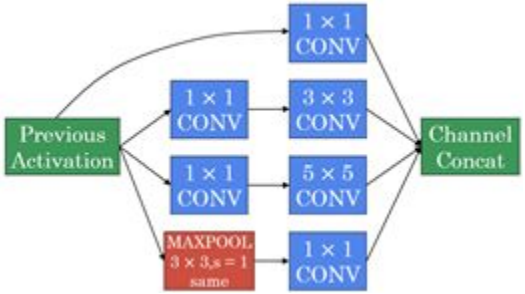
Using 1x1 convolution

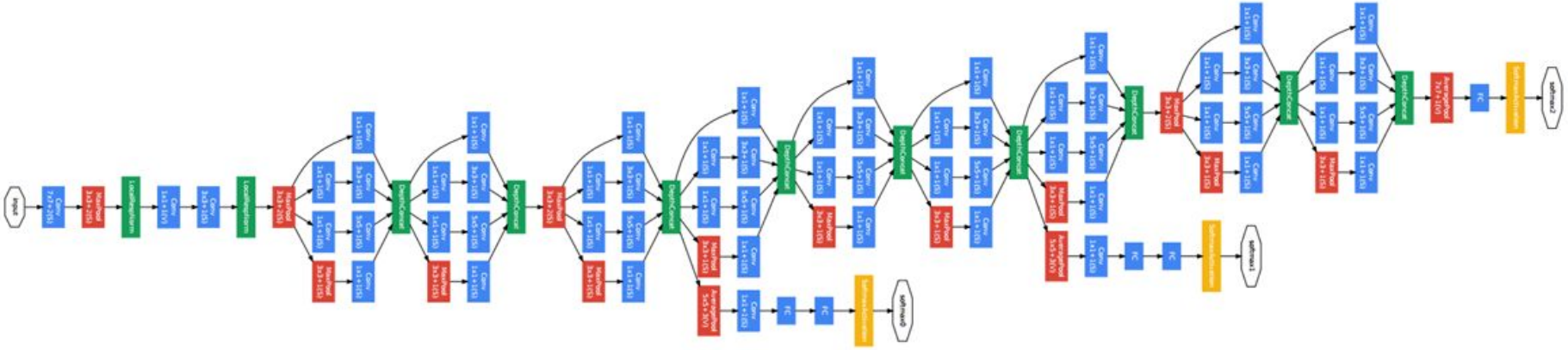


Inception Module



Inception network

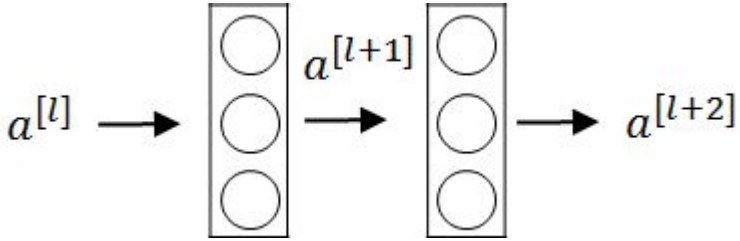




[Szegedy et al., 2014, Going Deeper with Convolutions]

Residual Networks (ResNets)

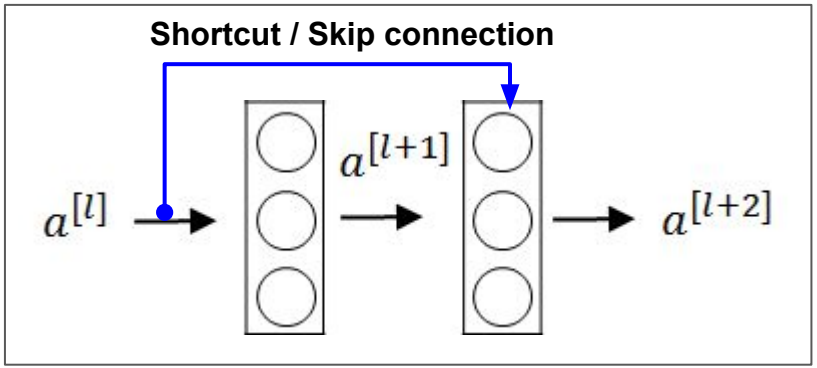
Residual Block



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

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

Residual Block



$$\mathbf{z}^{[l+1]} = W^{[l+1]}\mathbf{a}^{[l]} + \mathbf{b}^{[l+1]}$$

$$\mathbf{a}^{[l+1]} = g(\mathbf{z}^{[l+1]})$$

$$\mathbf{z}^{[l+2]} = W^{[l+2]}\mathbf{a}^{[l+1]} + \mathbf{b}^{[l+2]}$$

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

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

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

ResNet