$6 \times 6 \times 3$

$a^{[0]}$

$x$

$z^{[1]} = W^{[1]} a^{[0]} + b^{[1]}$

$a^{[1]} = g(z^{[1]})$

2 filters

$3 \times 3 \times 3$

$3 \times 3 \times 3$

"$W^{[1]}$"

"$W^{[1]} a^{[0]}$"

Relu

$+ b_1$

$4 \times 4$

Relu

$+ b_2$

$4 \times 4$

R

$Z^{[1]}$

$4 \times 4$

$4 \times 4$

$6 \times 6 \times 3$   $a^{[0]}$   $4 \times 4 \times 2$   $a^{[1]}$
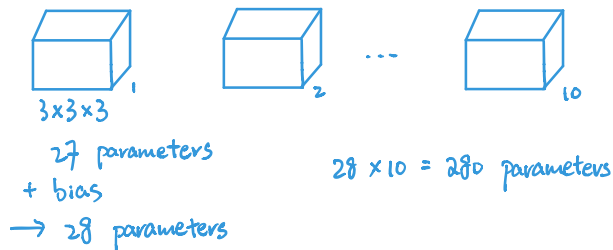
$4 \times 4 \times 2$

$\leftarrow a^{[1]}$

# Number of parameters in one layer

If you have 10 filters that are $3 \times 3 \times 3$ in one layer of a NN. How many parameters does that layer have?



$3 \times 3 \times 3$

1   2   ...   10

27 parameters
+ bias
$\rightarrow$ 28 parameters

$28 \times 10 = 280$ parameters

# Summary of notation

If layer $\ell$ is a convolution layer

$f^{[\ell]}$ = filter size

$p^{[\ell]}$ = padding

$S^{[\ell]}$ = Stride

$n_c^{[\ell]}$ = number of filters

Each filter is: $f^{[\ell]} \times f^{[\ell]} \times \boxed{n_c^{[\ell-1]}}$

Activation: $a^{[\ell]} \longrightarrow n_H^{[\ell]} \times n_W^{[\ell]} \times n_c^{[\ell]}$

Weights: $f^{[\ell]} \times f^{[\ell]} \times n_c^{[\ell-1]} \times n_c^{[\ell]}$

bias: $n_c^{[\ell]} - (1,1,1,n_c^{[\ell]})$ ← # of filters in layer $\ell$

Input: $n_H^{[\ell-1]} \times n_W^{[\ell-1]} \times \boxed{n_c^{[\ell-1]}}$

Output: $n_H^{[\ell]} \times n_W^{[\ell]} \times n_c^{[\ell]}$

$$n_{\substack{H \\ W}}^{[\ell]} = \left\lfloor \frac{n_{\substack{H \\ W}}^{[\ell-1]} + 2p^{[\ell]} - f^{[\ell]}}{S^{[\ell]}} + 1 \right\rfloor$$

$A^{[\ell]} \longrightarrow m \times n_H^{[\ell]} \times n_W^{[\ell]} \times n_c^{[\ell]}$