Why Convolutions? *CORRECTION*

Starting around 2:15 minute, the number of parameters should have been:

(5*5*3+1)*6 = 456

This is based on the equation:

$(f^{[l]} \times f^{[l]} \times n_c^{[l-1]} + 1) \times n_c^{[l]}$.

$f^{[l]}$ is the filter height (and width).

$n_c^{[l-1]}$ is the number of channels in the previous layer.

$n_c^{[l]}$ is the number of channels in the current layer.

The "1" is the bias term.

(It was (5*5+1)*6=156 in the video.)