Suppose we have an input image **X**, with a width of **w** and height **h** and the filter of size PxQ, then the convolution operation can be represented as follows:

The equation basically represents how the output, , (that is the element of the ith row and jth column of the output matrix) using the convolution operation.

Once the convolution operation is performed, we feed the result to a feed forward network, f, and predict the output, :