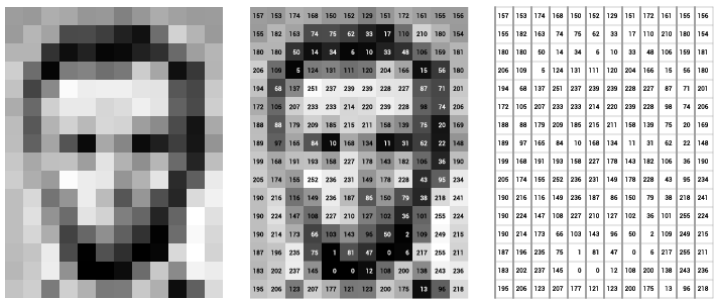
Deep Learning

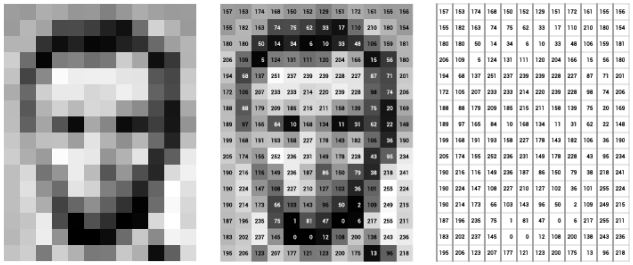
CNN: Input Channel



Learning goals

- Input Channel for Gray Images
- Input Channels for Colored Images
- Output Dimension Calculation

INPUT CHANNEL



- An image consists of the smallest indivisible segments called pixels and every pixel has a strength often known as the pixel intensity.
- A grayscale image has a single input channel and the value of each pixel represents the amount of light.
- A grayscale value can lie between 0 to 255, where 0 value corresponds to black and 255 to white.

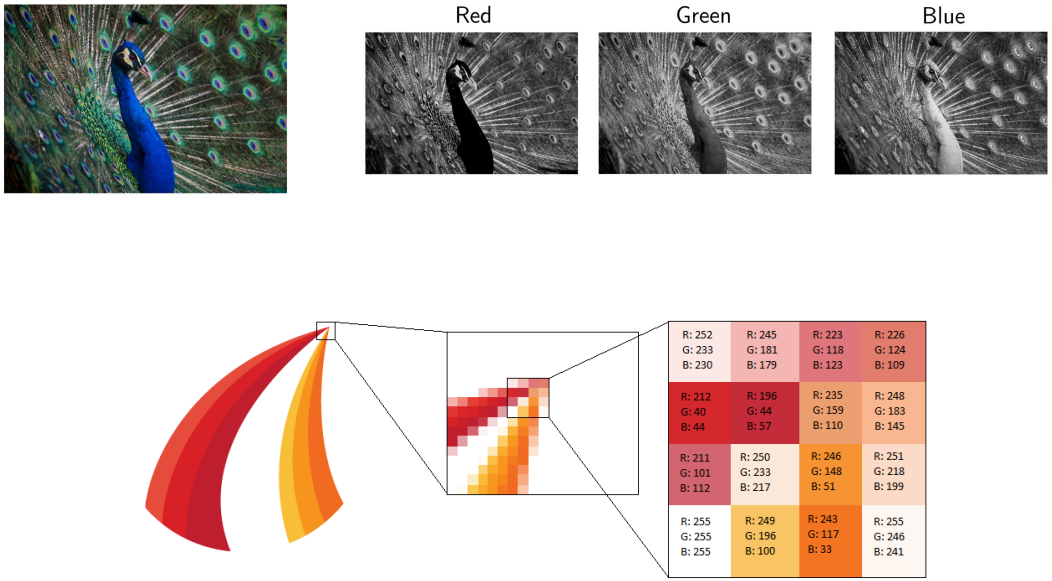


Figure: Image source: Computer Vision Primer: How AI Sees An Imag eKishan Maladkar's Blog)

- A colored digital image usually comes with three color channels, i.e. the Red-Green-Blue channels or RGB values.
- Each pixel can be represented by a vector of three numbers (each ranging from 0 to 255) for the three primary color channels.

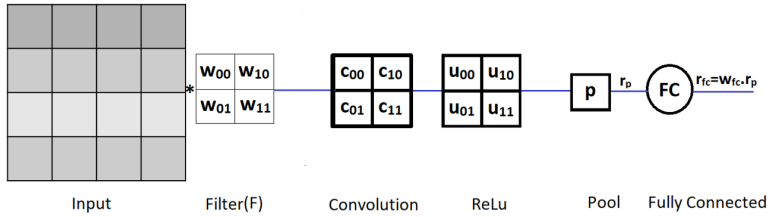


Figure: A CNN taking a grayscale image as input.

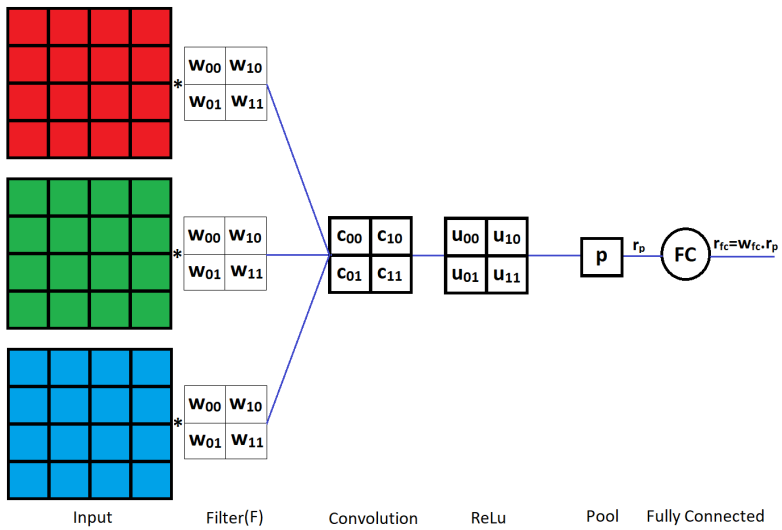
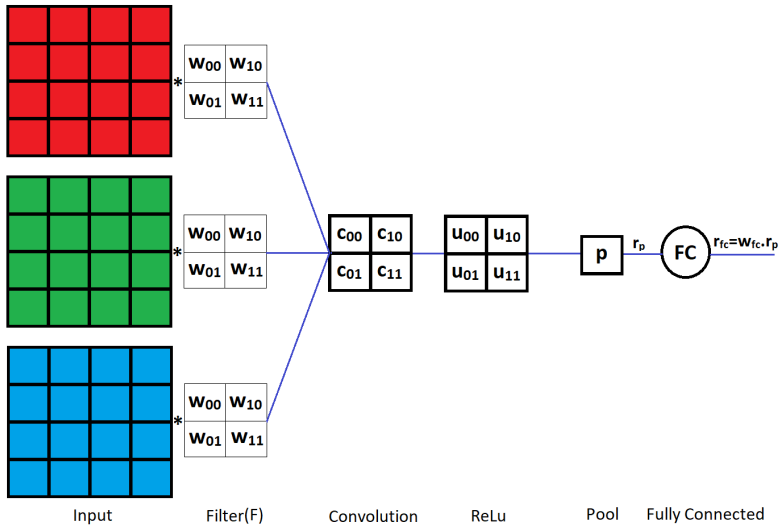


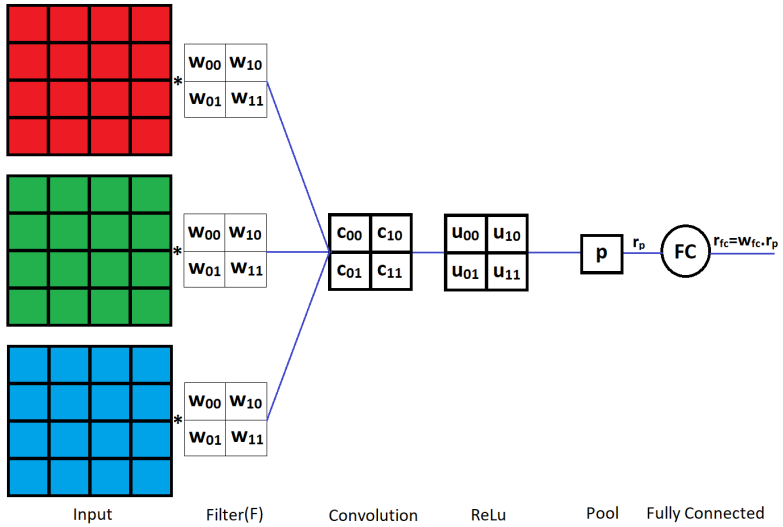
Figure: A CNN processing a colored images where each of the color spectrums serve as input. (Source: Chaitanya Belwal's Blog)



In this CNN:

- there are 3 input channels, represented as 4x4 input matrices,
- one 2x2 filter (also known as kernel),
- a single ReLu layer,
- a single pooling layer (which applies the MaxPool function),
- and a single fully connected (FC) layer.

- The elements of the filter matrix are equivalent to the unit weights in a standard NN and will be updated during the backpropagation phase.
- Assuming a stride of 2 with no padding, the output size of the convolution layer is determined by the following equation:
- $O = \frac{I-K+2.P}{S} + 1$ where:
  - O: is the dimension (rows and columns) of the output square matrix,
  - I: is the dimension (rows and columns) of the input square matrix,
  - K: is the dimension (rows and columns) of the filter (kernel) square matrix,
  - P: is the number of pixels (cells) of padding added to each side of the input,
  - S: is the stride, or the number of cells skipped each time the kernel is slided.



Inserting the values shown in the figure into the equation yields

$$O = \frac{I - K + 2.P}{S} + 1 = \frac{(4 - 2 + 2.0)}{2} + 1 \tag{1}$$

$$= 2. \tag{2}$$