Introduction to Deep Learning

Chapter 4: CNN: Input Channel

Bernd Bischl

Department of Statistics – LMU Munich

WS 2021/2022

INPUT CHANNEL

- An image consists of the smallest indivisible segments called pixels and every pixel has a strength often known as the pixel intensity. Whenever we study a digital image, it usually comes with three color channels, i.e. the Red-Green-Blue channels, popularly known as the RGB values.

- In colored images, each pixel can be represented by a vector of three numbers (each ranging from 0 to 255) for the three primary color channels: red, green, and blue.

- A grayscale image has a single input channel and value of each pixel represents only an amount of light;

- Note a grayscale value can lie between 0 to 255, 0 signifies black and 255 signifies white.

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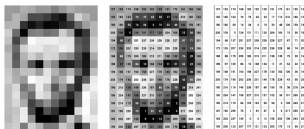
Figure: Image source: Computer Vision Primer: How AI Sees An Imag eKishan Maladkar's Blog)
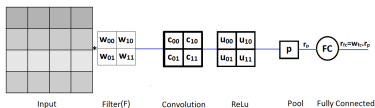
INPUT CHANNEL
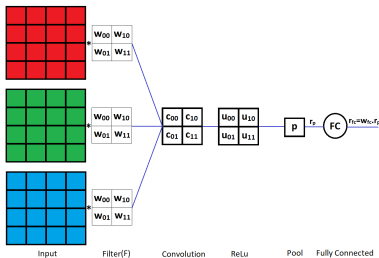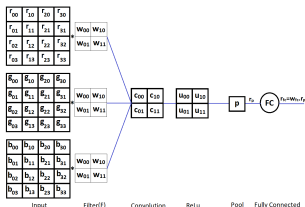

Figure: CNNs takes grayscale image as input.

Figure: CNNs use colored images where each of the Red, Green and Blue (RGB) color spectrums serve as input. (source: Chaitanya Belwal's Blog)

In this CNN:

- there are 3 input channel, with the size of 4x4 as an 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.

INPUT CHANNEL

- 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 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 new(convolution) 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,
  - S: is the stride, or the number of cells skipped each time the kernel is slided.

INPUT CHANNEL
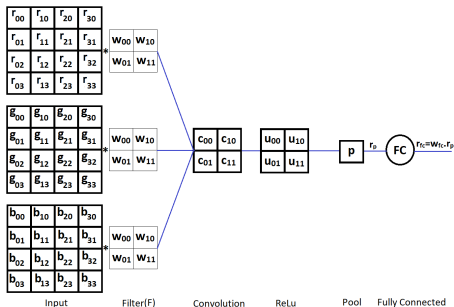
Putting these values in the equation,

$$O = \frac{(4 - 2 + 2.0)}{2} + 1 \tag{1}$$
$$= 2 \tag{2}$$