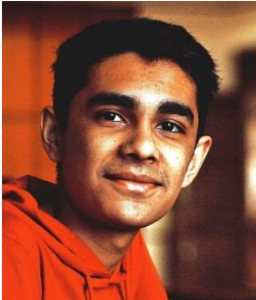


## CNN Model To Detect An Image



CNN



Human Face



CNN



Not a Human  
Face

The different layers in CNN are:

- Convolution
- ReLU layer
- Pooling
- Fully connected layer (ANN)

A computer understands an image using numbers at each pixels. Consider a grey scale image that has a value 1 for black pixel and 0 for white pixel. The input image is first passed through the convolution layer. CNN compares the image piece by piece i.e, looks for a particular feature. In this case it is a human face. It compares it with a feature

detector which we get after training of a dataset. The values in the image is multiplied with the corresponding values in feature detector. The average of this is taken and is called as filtered value. To keep track of where the feature was, we create a map and put the value of the filter at that place.

This filtered image is passed through the ReLU layer where an activation function known as Rectified Linear Unit is applied. The result of this is that we get a matrix that has values greater than zero if the condition is satisfied else zero.

The feature map is then maxpooled to get a pooled feature map. A pooled feature map is a shrunked matrix in which every value is found by taking a window size and then finding the maximum value in it. This value is placed in the corresponding position in the pooled feature map.

The pooled feature map is then linearly stacked (Flattening) and feed into ANN network where the actual prediction of the image takes place. The ouput we get is between 0 and 1. 1 implies that the image is a human face.