





**Aim**: To implement Convolutional Neural Network.

# Theory:

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. Convolutional neural networks are distinguished from other neural networks by their superior performance with image, speech, or audio signal inputs. They have three main types of layers as shown in figure 1, which are:

1. Convolutional layer
2. Pooling layer
3. Fully connected (FC) layer

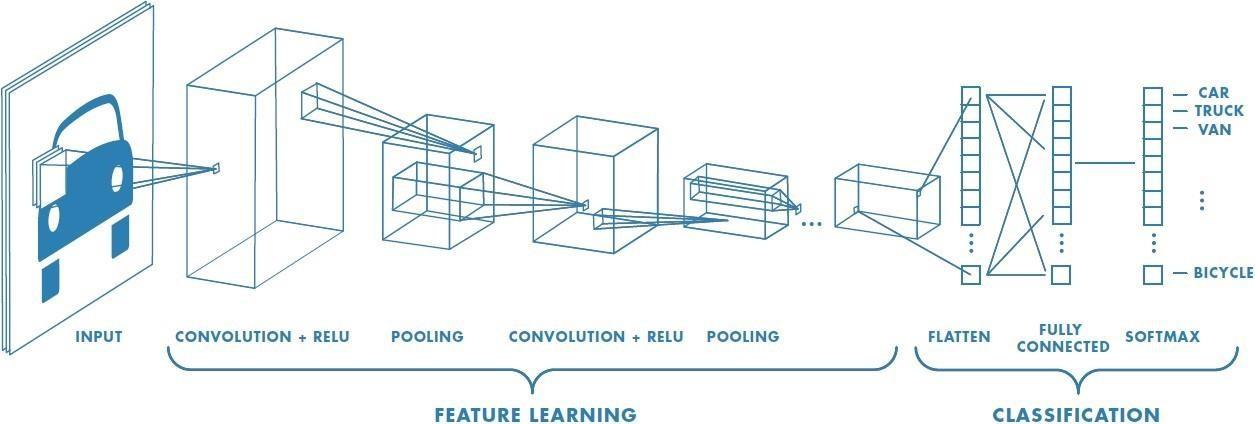


Figure 1: Architecture of CNN

The convolutional layer is the first layer of a convolutional network. While convolutional layers can be followed by additional convolutional layers or pooling layers, the fully connected layer is the final layer. With each layer, the CNN increases in its complexity, identifying greater portions of the image. Earlier layers focus on simple features, such as colors and edges. As the image data progresses through the layers of the CNN, it starts to recognize larger elements or shapes of the object until it finally identifies the intended object.

# Convolutional Layer

The convolutional layer is the core building block of a CNN, and it is where the majority of computation occurs. It requires a few components, which are input data, a filter, and a feature map. Let’s assume that the input will be a color image, which is made up of a matrix of pixels in 3D. This means that the input will have three dimensions—a height, width, and depth— which correspond to RGB in an image. We also have a feature detector, also known as a kernel or a filter, which will move across the receptive fields of the image, checking if the feature is present. This process is known as a convolution.

The feature detector is a two-dimensional (2-D) array of weights, which represents part of the image. While they can vary in size, the filter size is typically a 3x3 matrix; this also determines the size of the receptive field. The filter is then applied to an area of the image, and a dot product is calculated between the input pixels and the filter. This dot product is then fed into an output array. Afterwards, the filter shifts by a stride, repeating the process until the kernel has swept across the entire image. The final output from the series of dot products from the input and the filter is known as a feature map, activation map, or a convolved feature.

# Pooling Layer

Pooling layers, also known as down sampling, conducts dimensionality reduction, reducing the number of parameters in the input. Similar to the convolutional layer, the pooling operation sweeps a filter across the entire input, but the difference is that this filter does not have any weights. Instead, the kernel applies an aggregation function to the values within the receptive field, populating the output array. There are two main types of pooling:

* + **Max pooling**: As the filter moves across the input, it selects the pixel with the maximum value to send to the output array. As an aside, this approach tends to be used more often compared to average pooling.
  + **Average pooling**: As the filter moves across the input, it calculates the average value within the receptive field to send to the output array.

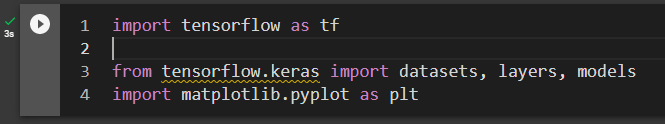
# Fully Connected Layer

The name of the full-connected layer aptly describes itself. As mentioned earlier, the pixel values of the input image are not directly connected to the output layer in partially connected layers. However, in the fully connected layer, each node in the output layer connects directly to a node in the previous layer. This layer performs the task of classification based on the features extracted through the previous layers and their different filters. While convolutional and pooling layers tend to use ReLu functions, FC layers usually leverage a softmax activation function to classify inputs appropriately, producing a probability from 0 to 1.

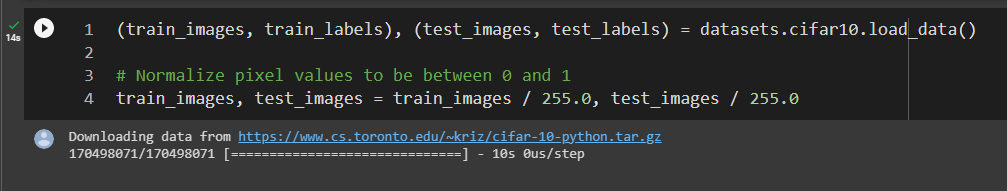


# Activity:

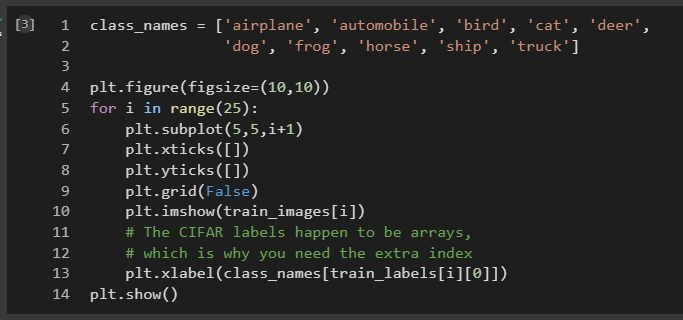
1. Import requisite libraries using Tensorflow and Keras.

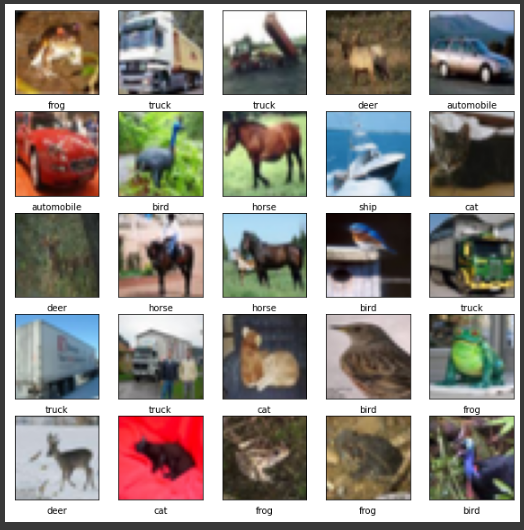


1. Load the selected dataset.

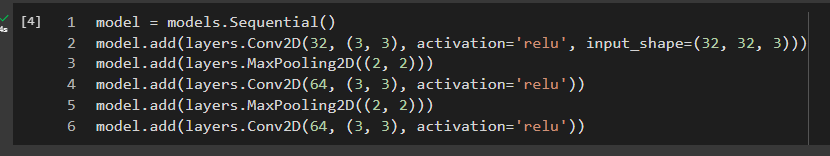


1. Visualize and display random images belonging to each class.

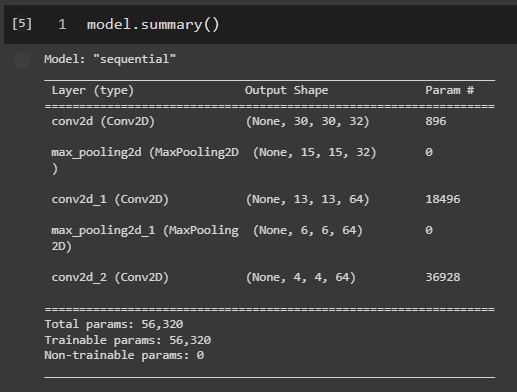


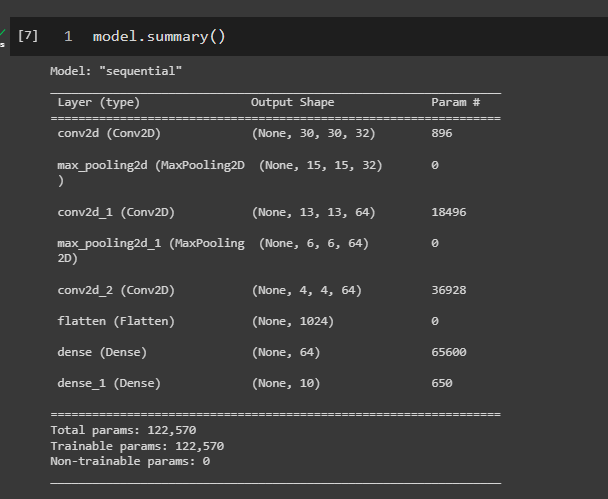


1. Develop the CNN model.

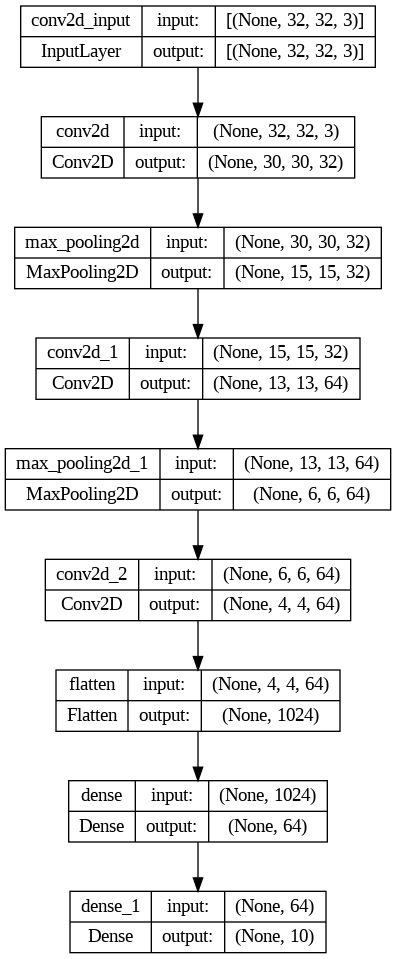


1. Print Model Summary and display architecture diagram.

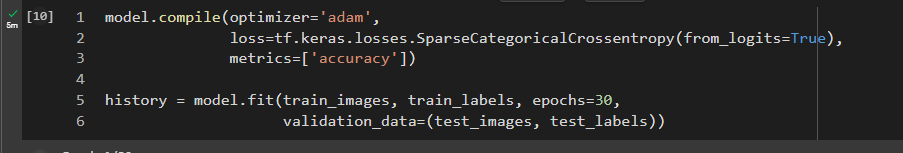


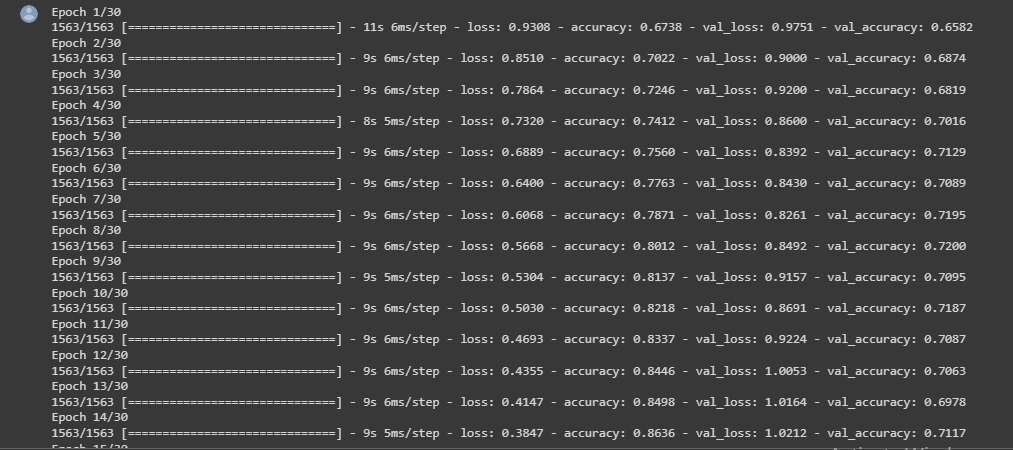






1. Compile and fit the model on train dataset.



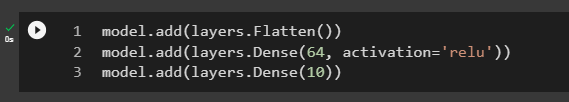


1. Calculate training and the cross-validation accuracy.

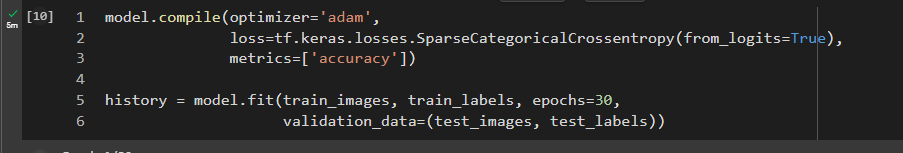


1. Redefine the model by using appropriate regularization techniques to prevent overfitting.

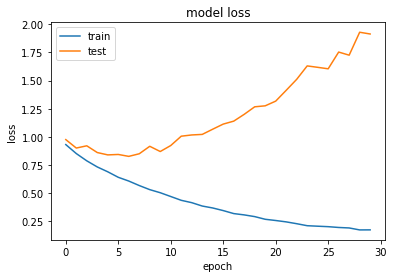
adding some extra dense layer

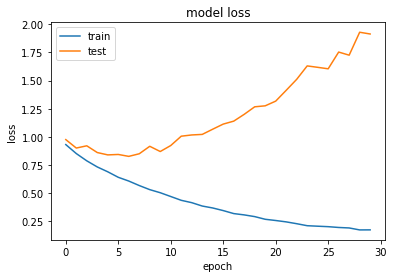


1. Fit the data on the regularized model.



1. Calculate and plot loss function and accuracy using suitable loss function.





1. Display classification Report for un-regularized CNN model.
2. Display classification Report for regularized CNN model.
3. Comment on output.



# CO3: Assimilate fundamentals of Convolutional Neural Network.



**Conclusion: Successfully understood how to implement Convolutional Neural Network.**



# Grade: AA / AB / BB / BC / CC / CD /DD

**Signature of faculty in-charge with date**



# References:

**Books/ Journals/ Websites:**