

1. The layers of the network are as follows:
  - a. **Input layer** : Initial layer of the network that contains the input data for training/ inference
  - b. **Convolution layer**: This layer performs the 2D convolution over the image matrix by using different filter sizes and user defined strides. The convolution layer is used to extract "features" from an image and different filters learn different weights during the model training that essentially refer to features of an image.
  - c. **Pooling layer**: The pooling layer often follows the convolution layer in a CNN. This layer helps to reduce the spatial complexity of the network by reducing the number of parameters and also helps alleviate the problem of overfitting.
  - d. **Fully connected**: The final fully connected layer uses a softmax function to compute prediction probabilities for the output classes.
2. Increasing the number of filters & reducing the kernel size for each convolution layer, increases the accuracy of the network.
3. Removing the pooling layers reduces the accuracy to 84%
4. The best accuracy achieved is 95%