

What is cNN?

Ans: convolution neural network is one kind of deep learning method. It is used for different kinds of classification problems like image classification, speech recognition etc. It looks similar to Artificial Neural Network. But the major difference between CNN and ANN is the connection between one layer to another layer. ANN is a fully connected neural network. But in CNN each neuron does not need to connect all the previous neurons.

How does CNN works?

Ans: As we can see in the diagram there are mainly two sections of the network. One is feature learning and another is classification. We will focus on the feature learning section. In this section we can see some keywords like Convolution, relu, pooling. We will explain these later. And the classification section is a fully connected network which is some kind of traditional neural network. In this section some fancy calculations happen behind the scene. One of the most important thing is back propagation. With this back propagation technique the network tracks those previous filters and tune so that it performs better in the next iteration.

Convolution layer?

Ans: we have to know that the computer sees the image as a matrix. Those are nothing but every individual pixel value of the image. If the image is black and white then the matrix will be two dimensional and if the image has RGB value that means color image then the matrix will be three dimensional. For simplicity we use a 2 dimensional image to explain the mechanism. In the diagram the input image is 5 by 5. In the convolution layer there will be multiple neurons. Those neurons are also called filter. A filter is also a matrix but smaller than the input matrix. In a single convo layer every filter dimension are same but values may differ. When we feed an image to one of this filter it will slide the filter matrix on image, do dot product between image pixel value and matrix value and sum all the value and place it in another matrix which is the output of this layer.

It is common that when a conv layer is used there will be a relu layer. Relu is an activation function and the full form is linear rectified unit. It gives nonlinearity to the network that the model can precisely separate different classes.

Pooling or maxpooling:

Ans: commonly maxpooling used after every conv layer. Conv layer map the features and maxpooling layer pool those features. Maxpooling always care about dominating feature. So it pools the maximum value from the previous layer output.

Why we use CNN?

Ans: At first we tried to achieve our target by machine learning method called KNN. But we did not get our desired accuracy. Because KNN works well on simple shaped images. After that we try to achieve our target by a deep learning method. There are lots of classification methods available right now. Some of them are mentioned below. Convolution neural network is one of the most popular deep learning method for image classification. And it outperforms previously established methods in terms of image classification. It is easy to use and understand. That's why we prefer CNN over other methods.

KNN working process:

These are some preprocess techniques our team member previously explain that. In KNN method there is no learning phase like CNN. So if we want to classify an image, we have to

calculate and compare it with all the images in the data set every time. If the data set is large then this method is not practical.

For testing the KNN method we use n fold cross validation technique. In our case the value of n is 5.

KNN result:

At maximum we get 63.5% accuracy in KNN method. This did not meet our target. That's why we move on to the CNN method.

CNN working process:

We use a little bit different technique to preprocess the data for CNN method. This was also described previously in the data preprocess section.

CNN is a learning based method. And filter number and filter size are very important in it. We design the model structure in such a way that we can tune filter number and filter size in every layer.

CNN result:

Tuning those parameters we generate 81 different learning models. And here is the summary of the top 10 best performing models. Among all the models, Model-44 performs best. So we use a large iteration for Model-44 and it gets very close to our target.