CONVOLUTION NEURAL NETWORK (CNN)

Dr. Jagriti Saini

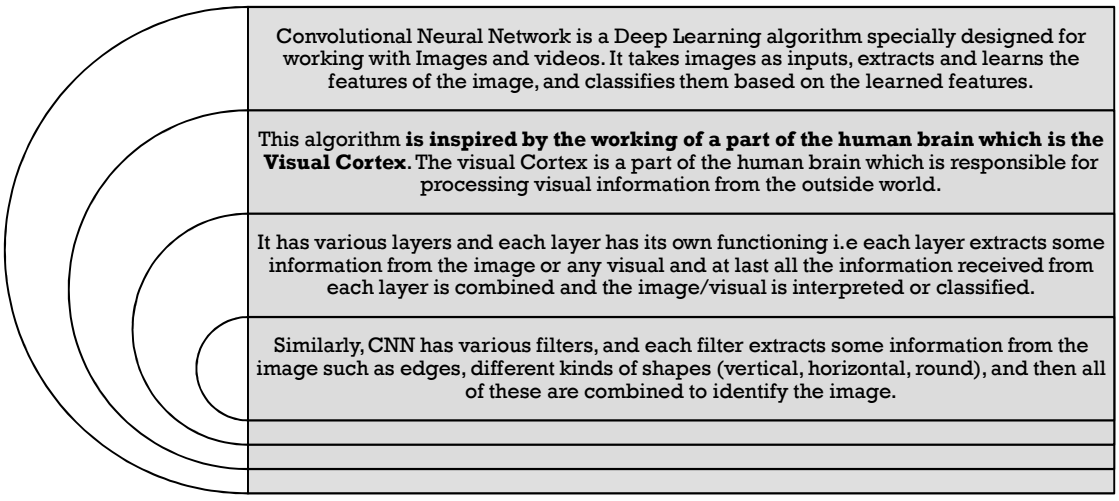
Founder & Owner

Eternal RESTEM

Website: www.eternalrestem.com

Email: jagritis1327@gmail.com

INTRODUCTION TO CNN



Convolutional Neural Network is a Deep Learning algorithm specially designed for working with Images and videos. It takes images as inputs, extracts and learns the features of the image, and classifies them based on the learned features.

This algorithm **is inspired by the working of a part of the human brain which is the Visual Cortex**. The visual Cortex is a part of the human brain which is responsible for processing visual information from the outside world.

It has various layers and each layer has its own functioning i.e each layer extracts some information from the image or any visual and at last all the information received from each layer is combined and the image/visual is interpreted or classified.

Similarly, CNN has various filters, and each filter extracts some information from the image such as edges, different kinds of shapes (vertical, horizontal, round), and then all of these are combined to identify the image.

INTRODUCTION TO CNN

Now, the **question here can be: Why** can't we use Artificial Neural Networks for the same purpose? This is because there are some disadvantages with ANN:

- It is too much computation for an ANN model to train large-size images and different types of image channels.
- The next disadvantage is that it is unable to capture all the information from an image whereas a CNN model can capture the spatial dependencies of the image.
- Another reason is that ANN is sensitive to the location of the object in the image i.e if the location or place of the same object changes, it will not be able to classify properly.

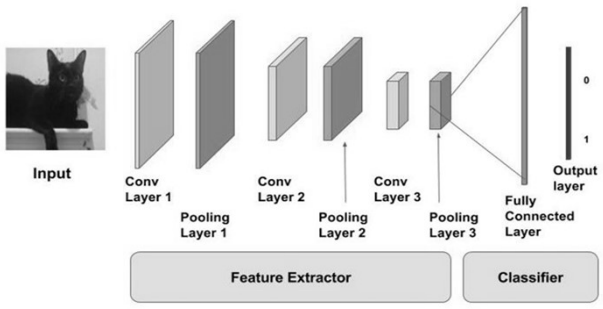
Why CNN?

- **Automatic Feature extraction** is ideal for image classification problems.
- In CNN all layers are not fully connected which reduces the amount of computation (which means fewer parameters to learn) unlike simple artificial neural networks.
- CNN's are invariant to the location of the object in the image and distortion in the scene.

COMPONENTS OF CNN

- The CNN model works in two steps: **Feature Extraction and Classification**
- **Feature Extraction** is a phase where various filters and layers are applied to the images to extract the information and features out of it and once it's done it is passed on to the next phase i.e **Classification** where they are classified based on the target variable of the problem.
- **A typical CNN model looks like this:**
  - Input layer
  - Convolution layer + Activation function
  - Pooling layer
  - Fully Connected Layer

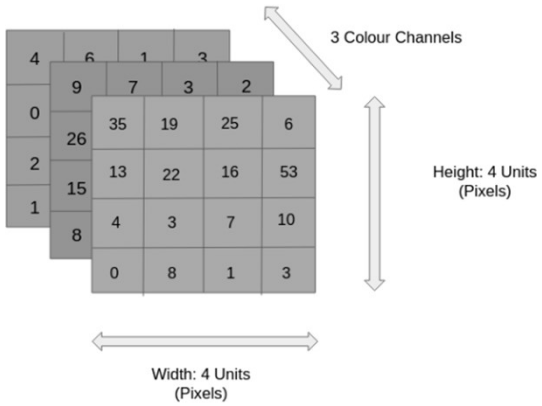
  Few Additional Layers:
  - Batch Normalization
  - Flatten Layer
  - Dense
  - Dropout



COMPONENTS OF CNN

Input layer

- As the name says, it's our input image and can be Grayscale or RGB.
- Every image is made up of pixels that range from 0 to 255. We need to normalize them i.e convert the range between 0 to 1 before passing it to the model.
- Here is the example of an input image of size 4*4 and has 3 channels i.e RGB and pixel values.



COMPONENTS OF CNN

Convolution Layer

- The convolution layer is the layer where the **filter is applied to our input image** to extract or detect its features.
- A filter is applied to the image multiple times and creates a feature map which helps in classifying the input image: normalized pixels.

0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1

6*6

*

1	0	-1
2	0	-2
1	0	-1

3*3

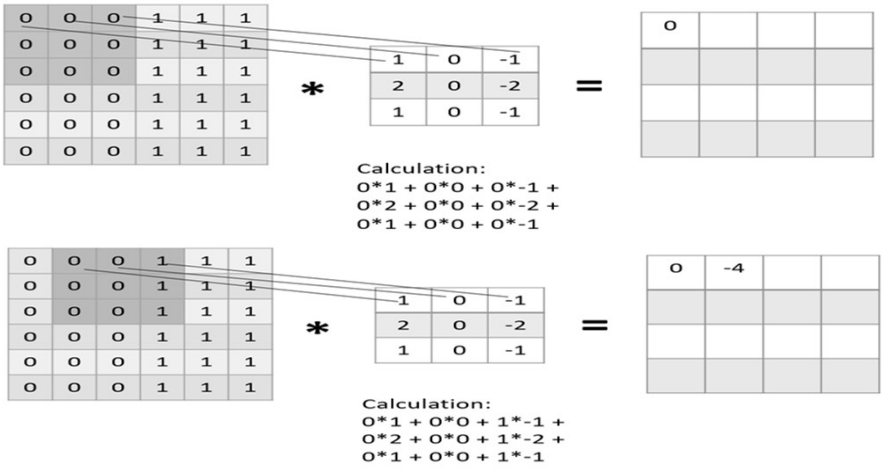
=

0	-4	-4	0
0	-4	-4	0
0	-4	-4	0
0	-4	-4	0

4*4

COMPONENTS OF CNN

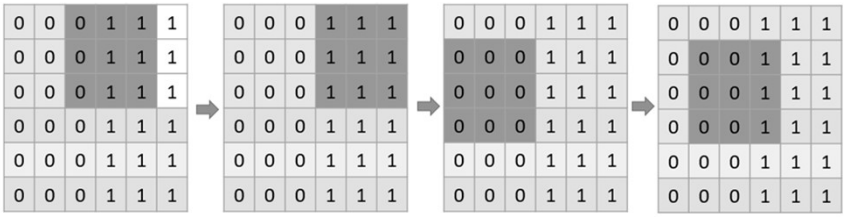
Convolution Layer



COMPONENTS OF CNN

Convolution Layer

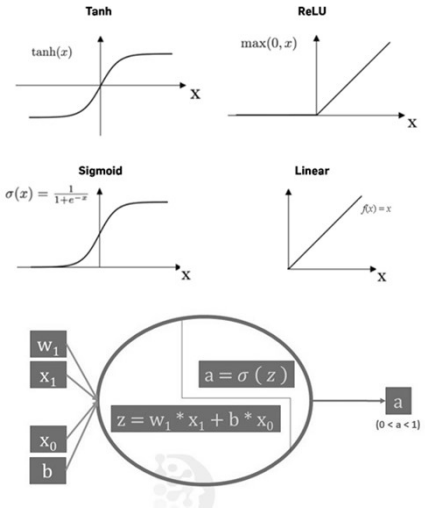
- Similarly, the filter passes over the entire image and we get our final **Feature Map**. Once we get the feature map, an activation function is applied to it for introducing nonlinearity.
- A point to note here is that the Feature map we get is smaller than the size of our image. As we increase the value of stride the size of the feature map decreases.



COMPONENTS OF CNN

Role of Activation Function

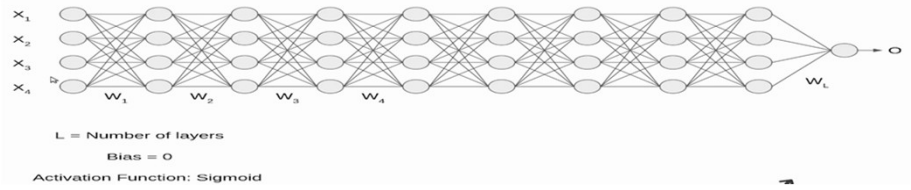
- The purpose of activation functions is mainly to add non-linearity to the network, which otherwise would be only a linear model.
- A convolutional layer by itself is linear exactly like the fully connected layer.
- The activation function calculates a weighted total and then adds bias to it to decide whether a neuron should be activated or not. The Activation Function's goal is to introduce non-linearity into a neuron's output.



COMPONENTS OF CNN

Batch Normalization Layer

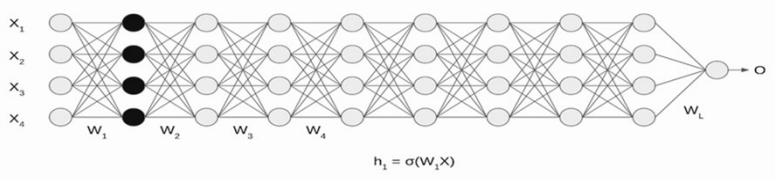
- Batch normalization is a process to make neural networks faster and more stable through adding extra layers in a deep neural network. The new layer performs the standardizing and normalizing operations on the input of a layer coming from a previous layer.
- But what is the reason behind the term "Batch" in batch normalization? A typical neural network is trained using a collected set of input data called **batch**. Similarly, the normalizing process in batch normalization takes place in batches, not as a single input.
- Let's understand this through an example, we have a deep neural network as shown in the following image.



COMPONENTS OF CNN

Batch Normalization Layer

- Initially, our inputs X1, X2, X3, X4 are in normalized form as they are coming from the pre-processing stage. When the input passes through the first layer, it transforms, as a sigmoid function applied over the dot product of input X and the weight matrix W.



- Similarly, this transformation will take place for the second layer and go till the last layer.
- Although, our input X was normalized with time the output will no longer be on the same scale. As the data go through multiple layers of the neural network and L activation functions are applied, it leads to an internal co-variate shift in the data.

COMPONENTS OF CNN

Pooling Layer

- The pooling layer is applied after the Convolutional layer and is used to reduce the dimensions of the feature map which helps in preserving the important information or features of the input image and reduces the computation time.
- Using pooling, a lower resolution version of input is created that still contains the large or important elements of the input image.
- The most common types of Pooling are Max Pooling and Average Pooling. While max pooling gives the most prominent feature in a particular patch of the feature map, average pooling gives the average of features present in a patch.

COMPONENTS OF CNN

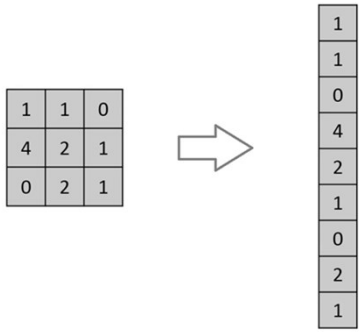
Pooling Layer

- The below figure shows how Max Pooling works. Using the Feature map which we got from the above example to apply Pooling. Here we are using a Pooling layer of size 2*2 with a stride of 2.
- The maximum value from each highlighted area is taken and a **new version of the input image is obtained which is of size 2*2 so after applying Pooling the dimension of the feature map has reduced.**

COMPONENTS OF CNN

Flattening Layer

- Basically flattening is taking matrix came from convolutional and pooling processes and turn it into one dimensional array.
- This is important because input of fully-connected layer -or let's say Artificial Neural Networks- consist of one dimensional array.



COMPONENTS OF CNN

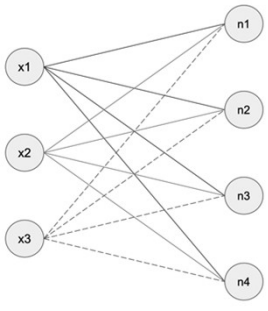
Fully Connected Layer

- Till now we have performed the Feature Extraction steps, now comes the Classification part. The Fully connected layer (as we have in ANN) is used for classifying the input image into a label.
- This layer connects the information extracted from the previous steps (i.e Convolution layer and Pooling layers) to the output layer and eventually classifies the input into the desired label.

COMPONENTS OF CNN

Dense Layer

- Dense layers are used when association can exist among any feature to any other feature in data point. Since between two layers of size n1 and n2 , there can n1∗n2 connections and these are referred to as Dense.
- The dense layer's neuron in a model receives output from every neuron of its preceding layer, where neurons of the dense layer perform matrix-vector multiplication.
- Matrix vector multiplication is a procedure where the row vector of the output from the preceding layers is equal to the column vector of the dense layer. The general rule of matrix-vector multiplication is that the row vector must have as many columns like the column vector.



COMPONENTS OF CNN

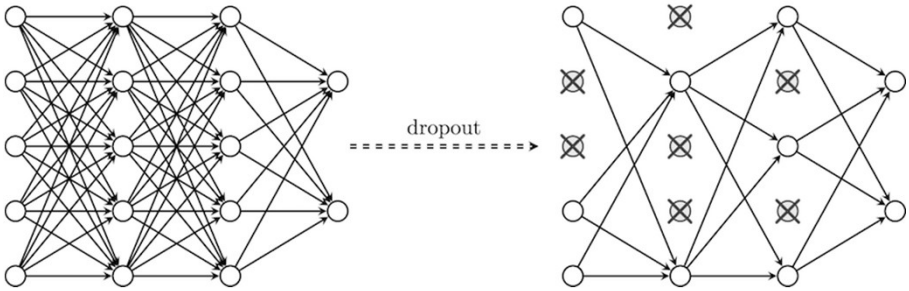
Dropout Layer

- Dropouts are the regularization technique that is used to prevent overfitting in the model. Dropouts are added to randomly switching some percentage of neurons of the network. When the neurons are switched off the incoming and outgoing connection to those neurons is also switched off.
- This is done to enhance the learning of the model. Dropouts are usually advised not to use after the convolution layers, they are mostly used after the dense layers of the network.

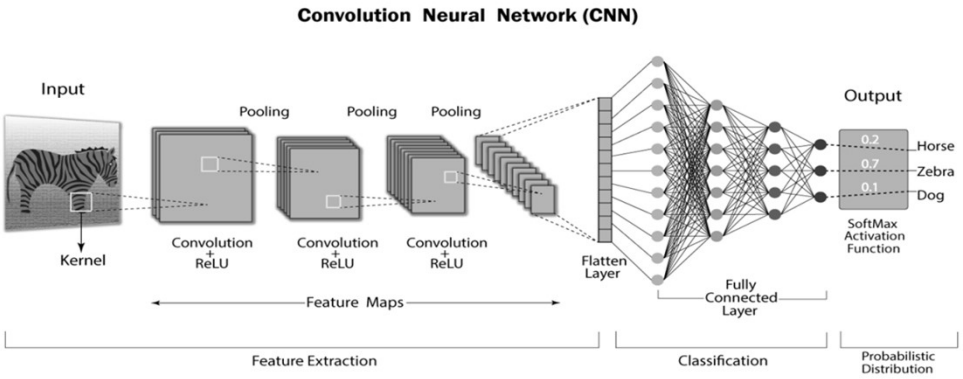
COMPONENTS OF CNN

Dropout Layer

- It is always good to only switch off the neurons to 50%. If we switched off more than 50% then there can be chances when the model leaning would be poor and the predictions will not be good.



COMPONENTS OF CNN



https://youtu.be/jDe5BAsT2-Y

THANK YOU