CNN → Convolutional Neural Network

3 layers:

Convolutional layer
Pooling layer
fully connected layer

Feature extraction :) separates and identifies the various features of the image.

convolutional layer :

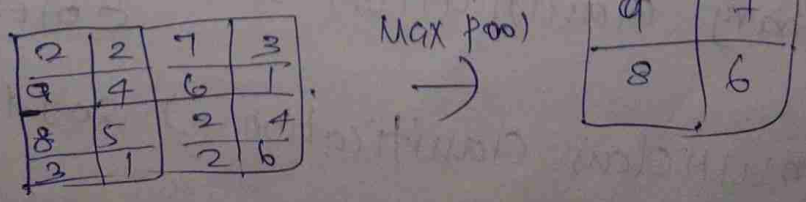
Convolution (mathametical operation) is performed between the input image and a filter of a particular size MXM. The output is feature map → which gives the information about the image.

Pooling layer :

It is used to reduce the dimensions of the feature map.

Max pooling :

It selects the maximum element from the region of the feature map covered by the filter.



Average pooling:

It computes average of the element in the region.

Global pooling:

It reduces each channel in the feature map to a single value.

Fully Connected Layer :- (FC)

The input image from the previous layers are flattened and fed to the FC Layer.

Dropout:

When all the features are connected to the FC layer, can cause overfitting. To overcome this problem a dropout layer is used.

Activation functions: (deciding factor)

It decides which information of the model should fire in the forward direction and which ones should not at the end of the network.

Binary classification -> sigmoid, softmax

multiclass classification -> softmax.

Maxpool layer -> for dimensionality reduction.

Padding ; - Striding :

VGG net :

VGGNet has 6 types.

The idea of the VGG architecture is stack the convolutional layers with increasing filters 512.
if layer1 has 16 filter than layer 2 must have 16 or more filters.

VGG 16 => 13 convolutional layers and
3 fully connected layers.
5 pooling layers.

All the hidden layers in the VGG network use ReLU.

AlexNet => 8 layers : 5 convolutional layer
3 fully connected layer

It uses ReLu.
It uses overlapping Max Pool (reduces overfitting & test ration error)

Advantage of convolution :- (In terms of feature)

i) Translation equivariance :
Some features are detected irrespective of Translation.

Advantage of pooling ; (In terms of pixel value)

i) Translation invariance :
Translation distance should not be high.

Padding:

The amount of pixels added to an image when it is processed by the kernel.

Adding padding to an image processed by a CNN allows for more accurate analysis of image.

Activation Function:

If we don't use activation function then the neural network is simpler and it will not be able to learn the complex patterns from the data.

sigmoid: range ⇒ 0 → 1

It is used for models where we have to predict the probability as an op.

sigmoid is for binary classification
softmax → multiclass classification

sigmoid looks at each raw output values separately so the op value not equal to the sum of 1.

softmax probabilities will always sum to 1.

Tanh: -1 to 1

RELU : Rectified Linear Unit [0 → ∞]

$f(z) \Rightarrow 0$ when z is less than 0.

$f(z) = z$ when $z >= 0$.

The negative values are turns into 0. So It affects the resulting graph by not mapping the negative values appropriately.

sigmoid and tanh create Vanishing gradiant problem. This problem is overcame in Relu. Bcaz the input gets large the slope doesn't saturate.

Relu → exploding gradiant.

All negative values are assign to 0. It creates a pblm called dying Relu.

over fitting - model is performing very well on the training data but is unable to predict the test data accurately.

To avoid this problem batch normalization is used. 2 step:-

input normalized

rescaling and offsetting.
(γ) (β)

Inception Resnet:
(Residual network)

InceptionNet: (27 layers)

Salient parts in the image can have extremely large variation in size. So choosing the right kernel size for the convolution operation becomes tough.

So in this filters have multiple size operates on the same level.

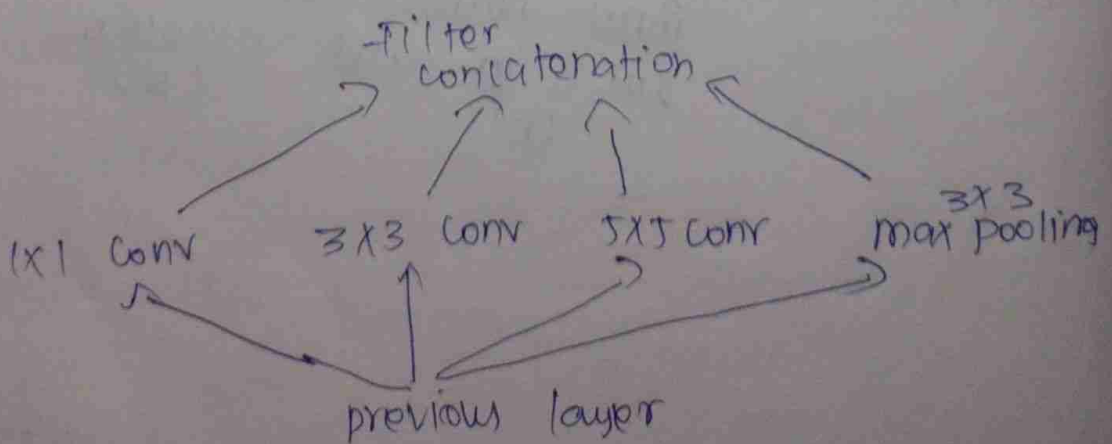
It performs convolution on an input, with 3 different filter size.
(1x1, 3x3, 5x5)

The op are concatenated and sent to the next inception module.

It takes the previous layer input and passes it into four different operations on parallel and concat the op from all the layers.

(1x1, 3x3, 5x5 convolution), (3x3 max pooling)

Computational cost is low.



Resnet Architecture :- (34 layer)

- To solve the vanishing/exploding gradiant, the Residual network is used.

skip connection → skip connections from few layers and conects directly to the op.

VGG 16 Implementation :

Sequential method → All the layers of the model will be arranged in sequence.

convolution + ReLu, max pooling
fully connected + ReLU, softmax

21 layers but 16 weight layers.

Input → (224 x 224)