**CNN**

CNN, Convnet or Convolutional neural network

It is a special kind of nn with Grid-like topology.

Biological inspiration with visual cortex (part of the brain responsible for object recognition).

MLP uses 1D data but cnn uses 2D data.

Hubel and Wiesel got the nobel prize for medicine in 1981. Here are the key finding of their cat experiment:

Some neurons in the visual cortex when presented line at specific angle

Our visual cortex is working with the certain parameters like V1 – Primary visual cortex which is also the edge part

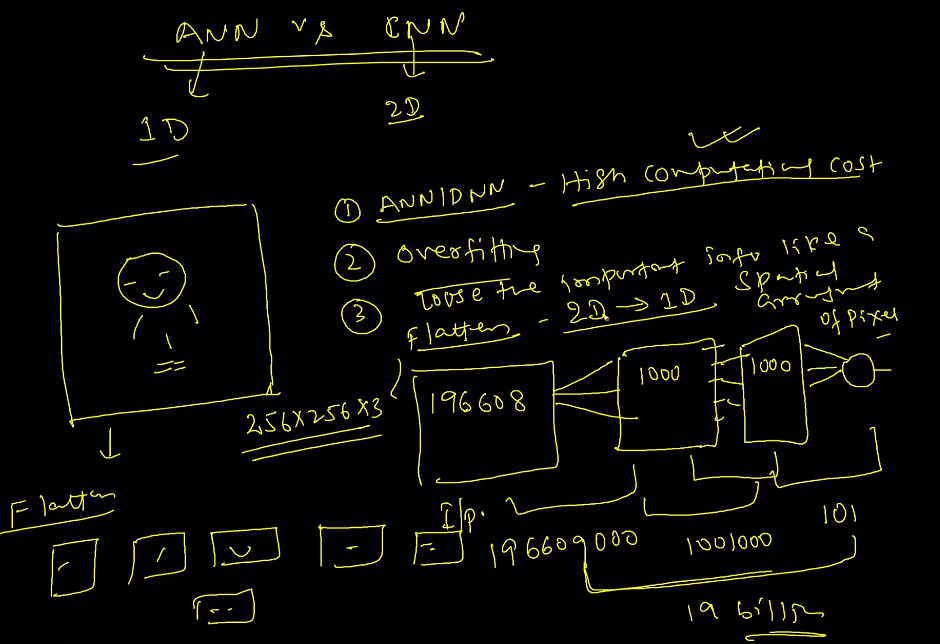
V2 – detect motion, depth, color, shape, faces.

Simple cell – smaller receptive field

Complex cell – bigger receptive field

Advantages of using CNN over ANN:

1. High computational cost
2. Overfitting
3. Loose the important information like spatial arrangement of pixels.



Certain characteristics of Convnet:-

* Conv Layer (Padding, Stride, Padding+Stride)
* Relu Layer (Max(0,x))
* Pooling layer (Min, Mean, Max, Global Pooling)
* Data Augmentation – Translation Invariance

Some popular Neural Networks in image processing:

* LeNet
* AlexNet
* VGG16/19
* ResNet
* Xception
* Transfer Learning
* MobileNet

Data Augmentation:-

Translation variance – the position of an object cannot change. It is significantly used in self-driving cars. It is a concept of computer vision.

But in CNN, we are just concerned with what that object is, and not its position or location.

Data Augmentation uses the concept of translation invariance as scaling, zooming in, zooming out, rotation, etc., hardly matters. Hence, CNN is a classification method unlike Computer vision, which also involves object detection, segmentation, recognition, etc.,

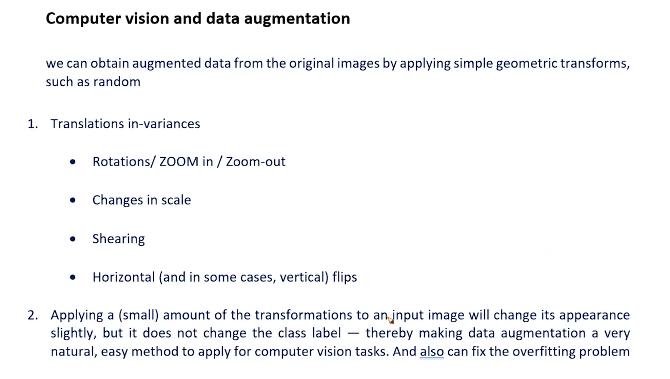
Why data augmentation is useful?

1. In-variance – shift invariance, rotation invariance, zoom invariance,

Shear invariance, Noise invariance, etc.,

1. Small Dataset – we can increase our data using DA.

Note:- Data Augmentation is very popular specifically for image dataset. The output will not change wrt to the input position.



CNN layers:-

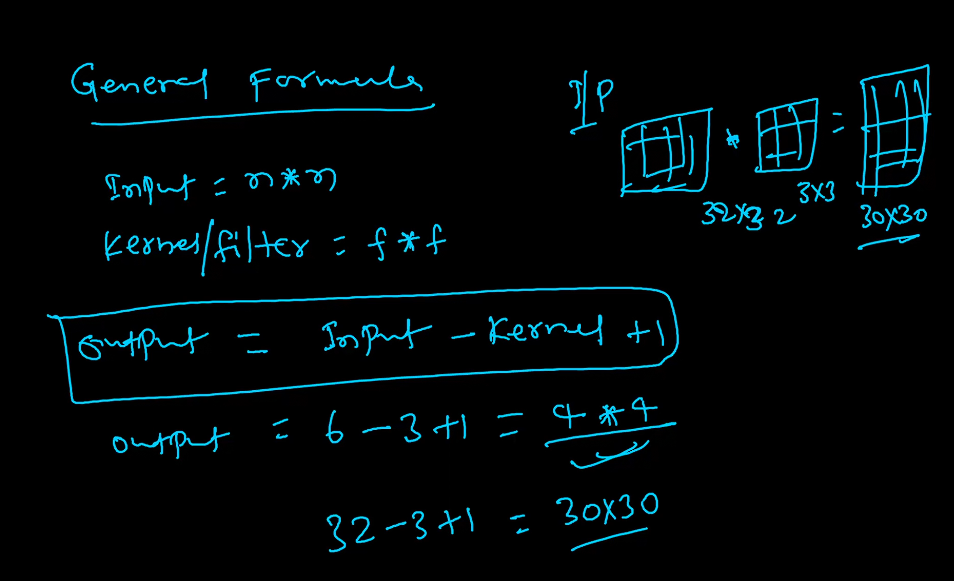
In CNN, we have 3 layers:

Convolutional layer

Relu layer

Pooling layer

Convolutional layer – there are 2 concepts in conv layers, Padding and Strides



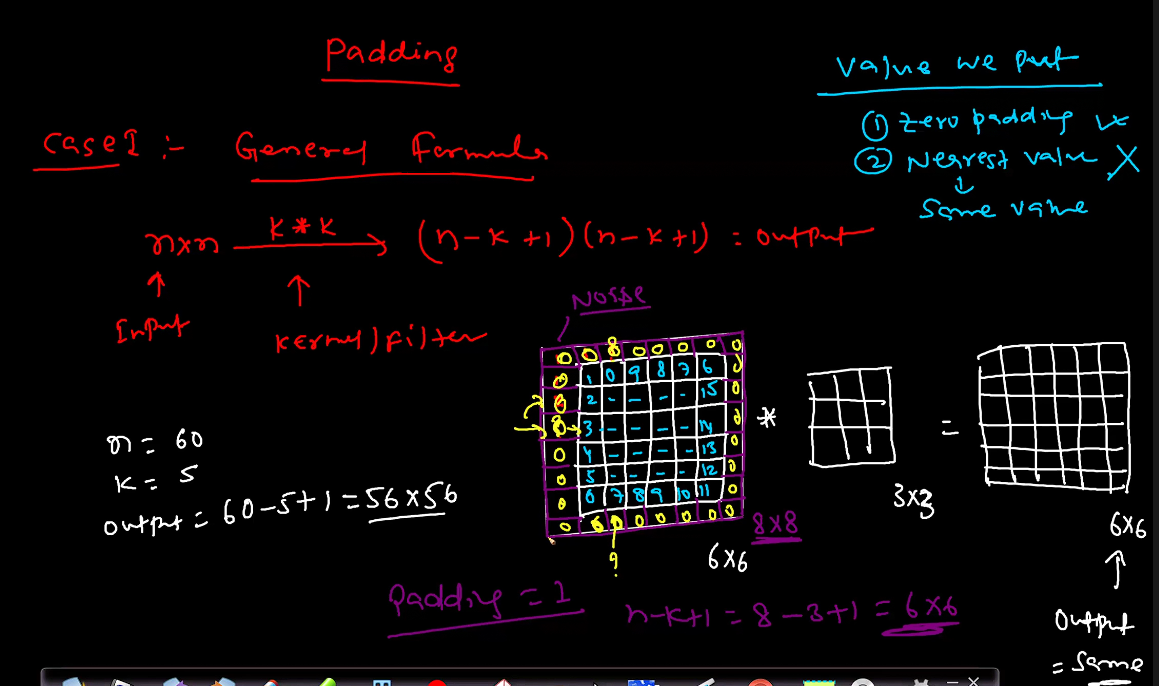
Padding:-

As per the general formula if my input size is 6x6 and my kernel or filter size is 3x3, the output that we shall get will be of size 4x4. But if we get a requirement that our output should also be of shape 6x6, we use a concept called Padding.

We add a layer around out input image to get a resulting image of size 8x8. Hence, we have achieved 8-3+1 = 6x6 output image. Here, padding = 1. We are adding noise to our data by introducing padding.

What values to put in the padding layer:-

We can put either zero (zero padding) or nearest value (same padding). Usually we go with zero padding.



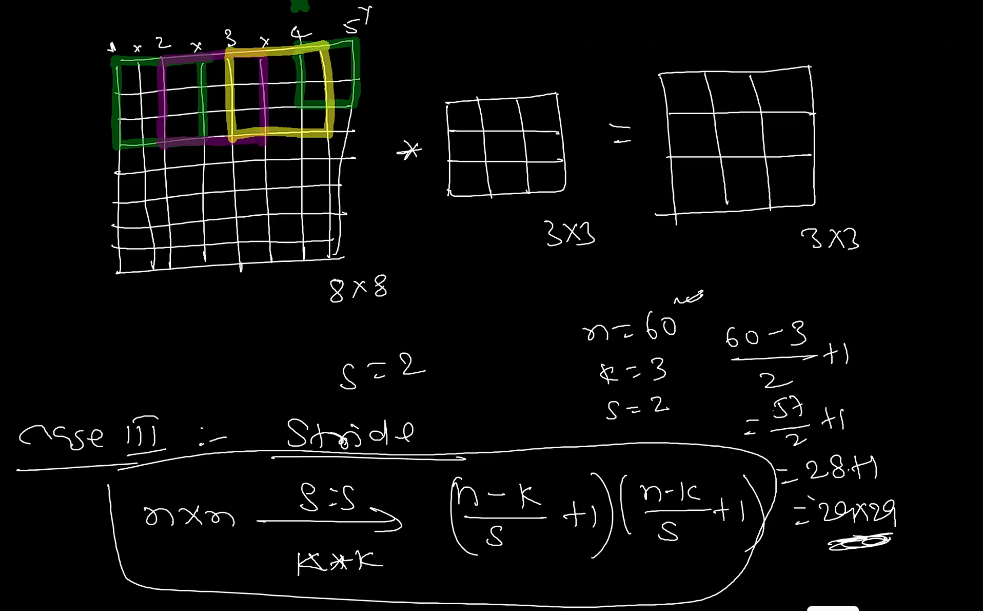
Stride:-

Stride means if we say shifted by 1, it means stride by 1

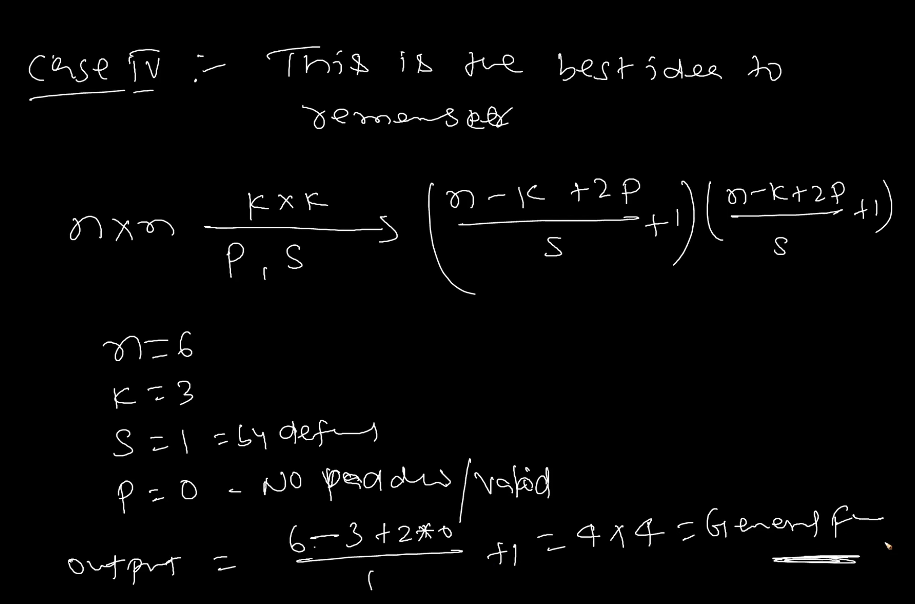
If we say shifted by 2, stride by 2.

If we say s = 2, it means shifting horizontally and vertically by 2 pixel.

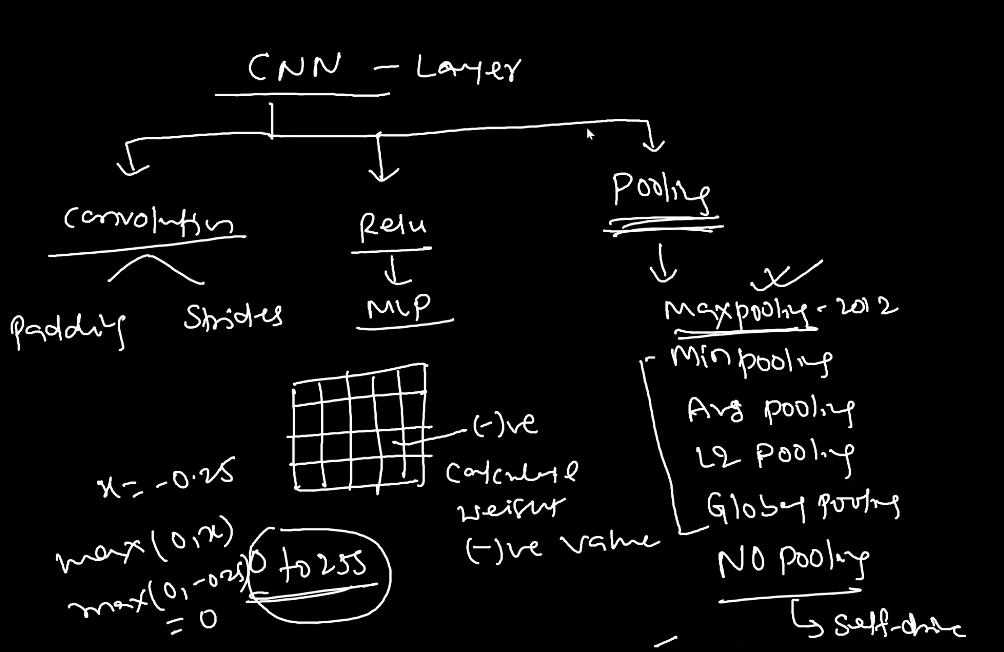
By default, stride = 1



Case 4:- This the main or the best idea to remember



Note:- We use Relu activation function to convert the negative weights (weights calculated can be negative as they are initialized by our algorithm like He-Initialization) to 0 as Relu is max(0, i) so max(0, -0.25) = 0. Since weights or pixel values(in case of CNN cannot be negative as they fall within the range of 0 to 255). Hence, we should only use Relu as the activation function in the case of CNN.



Now, comes the third layer, Pooling

We generally use Maxpooling which came in 2012. Its parameters comes under non-trainable params just like BatchNormzalization since we are just doing aggregation and no process like Back Propagation.

Advanatges:-

1. Reduces the size of input image so that the number of parameters can be controlled.
2. Translation Invariance (Makes the model less sensitive to the exact location of features).
3. Enhanced features(only in case of Maxpooling and not mean pooling etc.,) since they take the maximum or the best value.
4. No need of training for max pooling, mean pooling, min pooling etc.,
5. Reduced Computation: Smaller feature maps mean fewer parameters and faster training.
6. Focus on Key Features: Helps highlight the most important patterns.
7. Robust to Variations: Makes the model less sensitive to small changes in the input, like rotation or translation.

We should always take size = (2,2), strides = 2, type = MaxPooling while building our model.

Formula for calculating the parameters in our CNN model:-

**Kernel\_width \* kernel\_height \* no. of channels \* no. of filters applied + no. of filters applied**

It is analogous to no. of input\_data \* no. of weights + no. of bias in DNN.