CNN

Application of computer vision

1. Image classification -assigning a label or category to an entire image
2. Object detection -identifying and locating multiple objects within an image
3. Object localization -locating the position of objects within an image
4. Object detection and classification- simultaneously identifying and categorizing multiple objects
5. Semantic segmentation
6. Instance segmentation
7. Image generation – creating new image
8. Image style transfer – transferring artistic style from one image to another while retaining content.
9. Image super resolution – increasing the resolution and quality of an image
10. Image captioning – generating textual description of images automatically

Notes

Filters – we use for detecting the specific patterns in the input data. Each filter learns to recognize specific feature of the image

Stride – how many pixels the filter moves at the time of convolution

Padding – adding extra border pixels around the image or feature map before applying filters. Since the size of the output after filter reduces padding is done to save the original size of the image

Formula for calculating the output size of the convolutional layer = **n-f+1**

Where as n= height or width of the image, f= height or width of the filter/kernel

If the image is n\*m having fiter of p\*c then the output of the convolution will be

**(N-p+1)\*(m-c+1)**

If we are using the padding the formula becomes **[(n-f+2p)/s]+1**

Location invariance- the network can recognize features regardless of where they appear in the input. Network become robust to changes in location

Max pooling – it down samples the feature maps produced by the convolutional layer. It help to reduce complexity, reducing overfitting and helps to focus on most impornant features in each region.

There is also average pooling and min pooling which extract the average and min value from the convolution layer

Use cases

Max pooling – for edges , textures to get such important features

Mean pooling – it retain more information about the distribution and localization of features in each region of the image

Min pooling – specifically used for task such as detecting unusual patterns in the data such as outlier values etc. it also used to extract features having the low intensity

In back propagation the filter values are updated such that it extract the proper features we need for our required task such as image recognition or image classification etc

Data augmentation is used to increase the diversity and quality of the training data by applying various transofmation techniques such as rotation, scaling, flipping, or adding noise, zoom