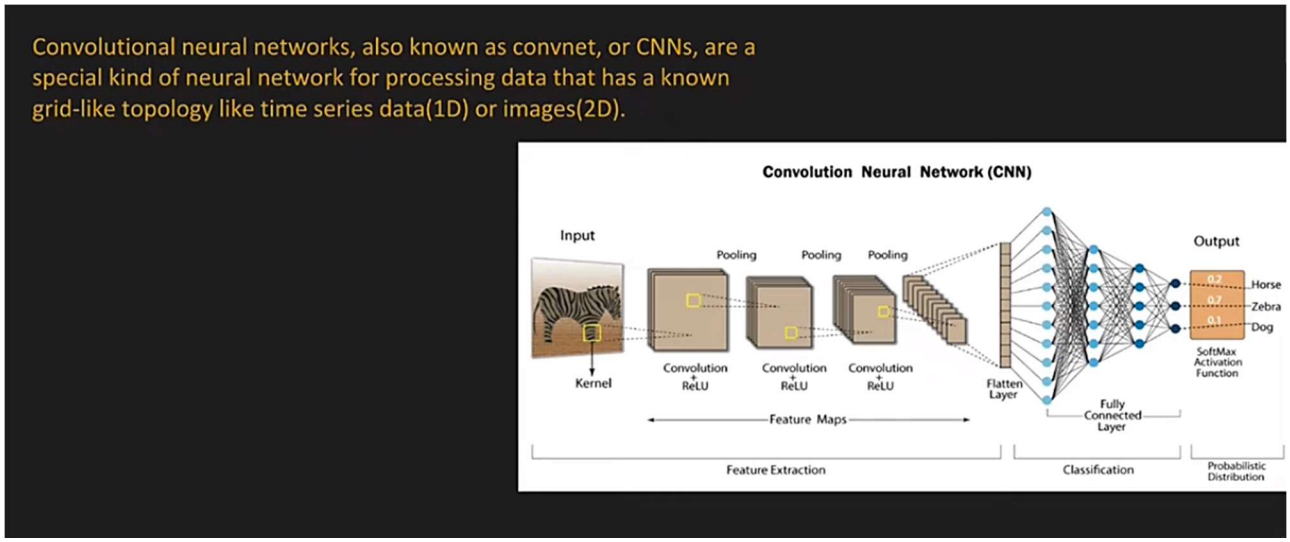
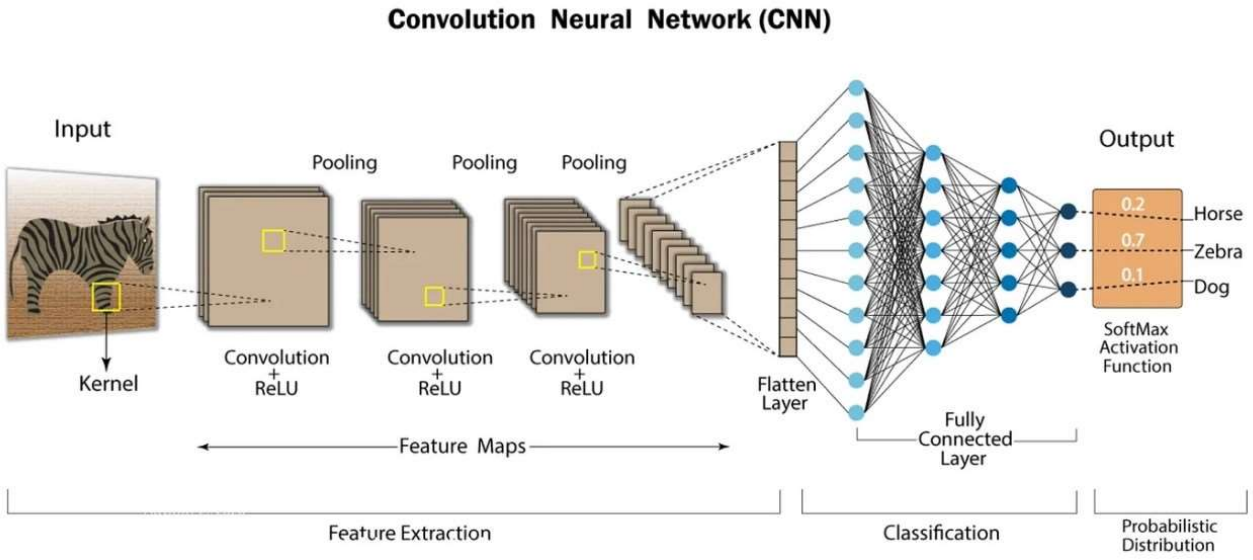
Conventional Neural Network (CNN)

- It is a filter from which image pass.
- Image is composed of pixels – pixels composed of grid base topologies
- Due to saturation – CNN can only see in images

What is CNN?

It is a special kind of neural network which deals with grid like topologies or time series data (1D or 2D)





Explanation

Feature Map

- First layer is Convolution layer (that is information carrier) – Then Pooling layer (after every conv a pooling layer add)

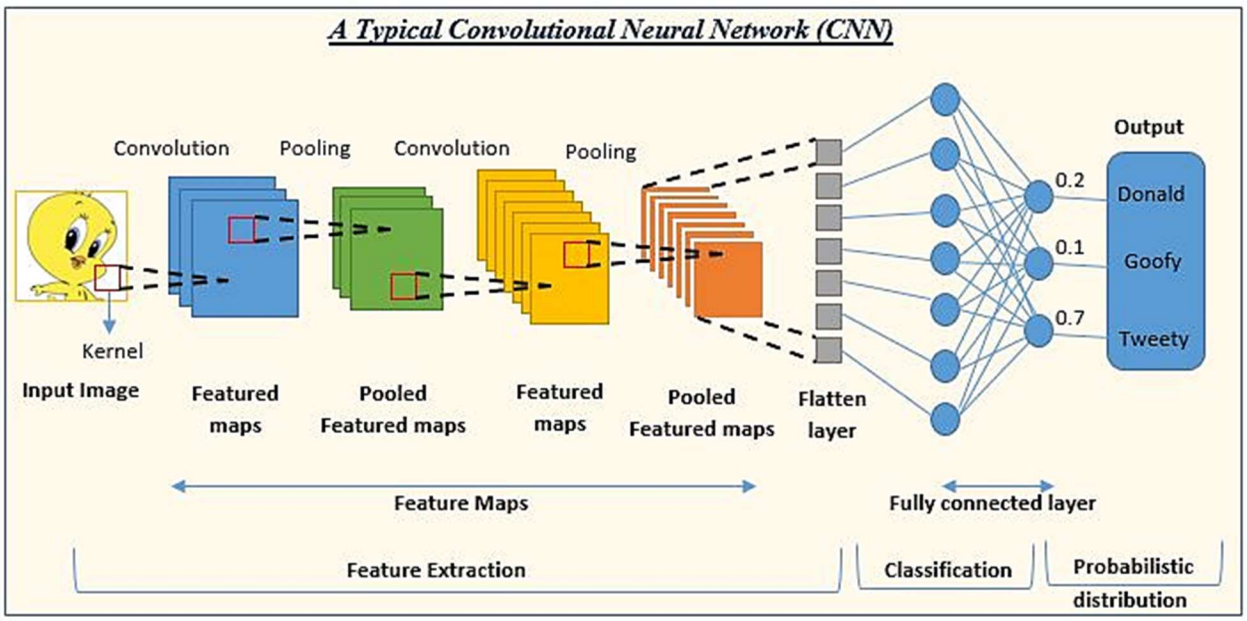
Fully connected layers

- Then Flatten layer – to transform data into vector / densor
- Then fully connected layer
- Then output layer

Note:

- Feature map means where filter is present (like CNN, RNN, GNN, Auto-encoders)
- Fully connected (like ANN)
- Feature extraction is automated (from input to flatten layer)
- Classification in fully connected layer
- Result in form of probabilistic distribution

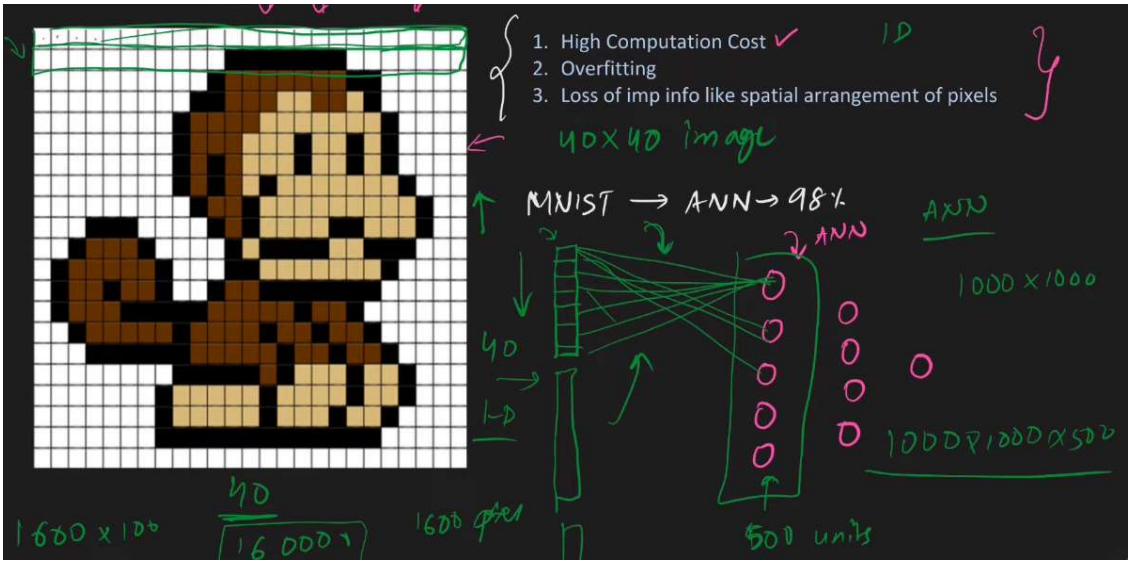
- probabilistic distribution – means by matching feature how much probability of an object match to original data or already trained image – means better feature extraction – depends on good data
- In image – every feature is important



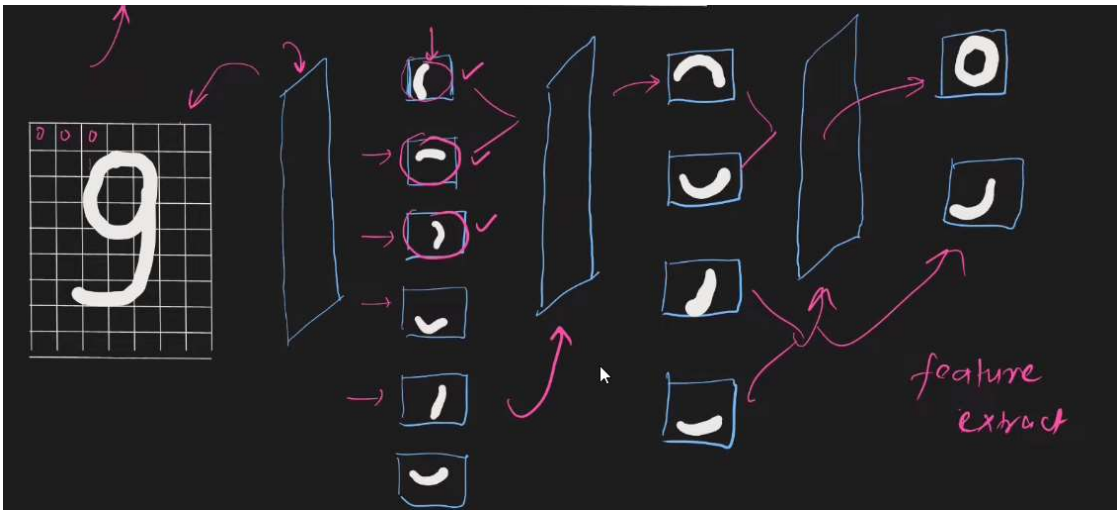
Why Not Use ANN?

- First reason – for images we do not use ANN because it cost large computational power
  E.g. 40x40 image cost huge vectors
- Second reason – Over fitting – it only learns feature but not its special arrangement of pixels
- Results are not good

That's why we have to more towards CNN

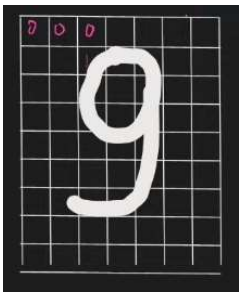


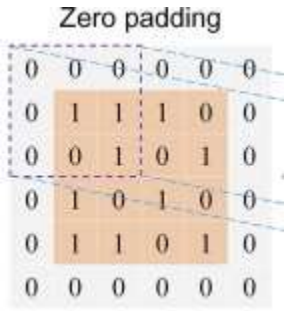
Intuition of CNN



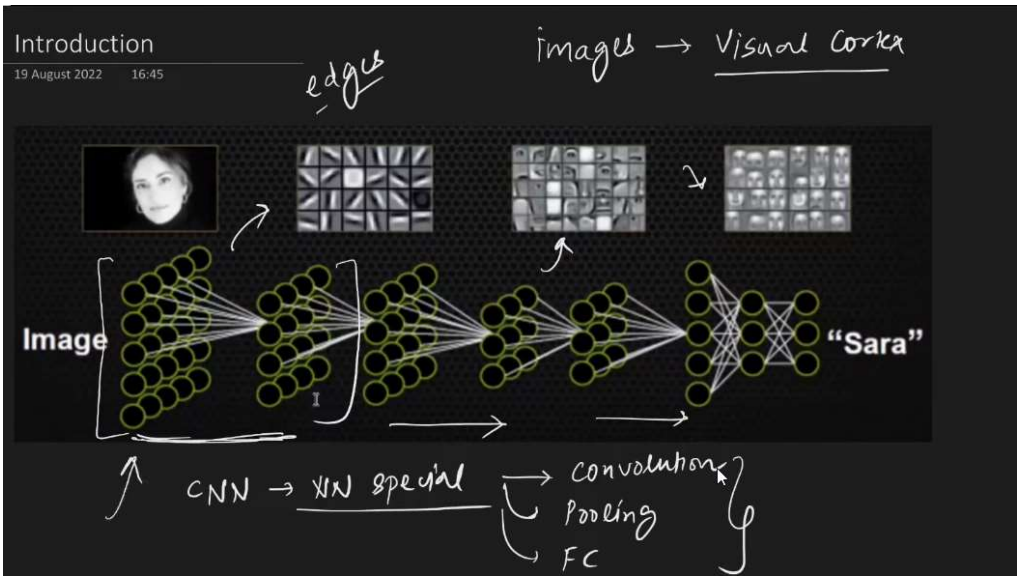
Explanation: Image9 pass through filter – then filter break it into parts

- Pass through first layer – it remember some parts
- Pass through second layer – it remember some parts and reduce its numbers
- Lastly, pass through last layer – it remembers some parts
- It stores information of every pixel in form of pattern like (where image is present it store as 1 and where there is no image it store as 0 vector)





- So first it goes through filter – some changes can be seen
- Then pass through pooling layer – we get summarized version
- Then goes through dense layer
- Then feature extract and flatten layer



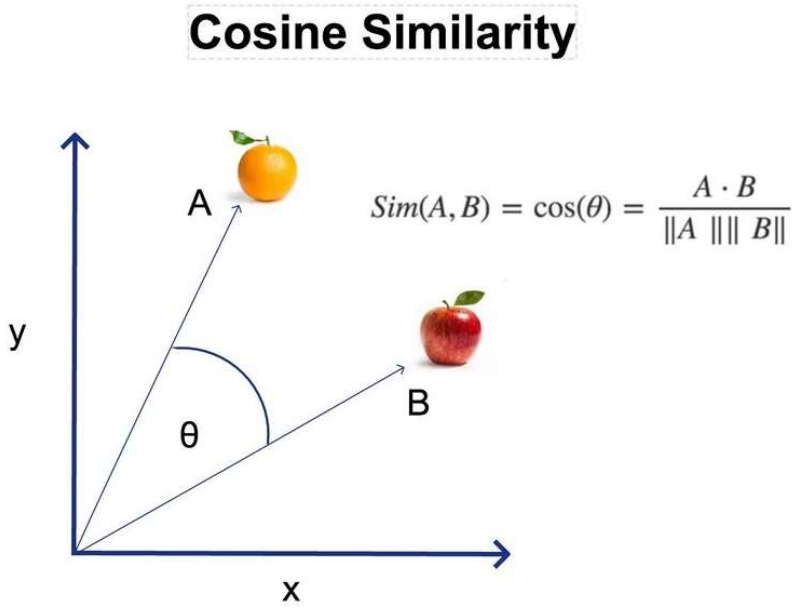
Explanation: In real image first – store edges – then combine edges – then image formed

- First layer – gets edges only
- Second layer – gets edges plus perceive information
- Lastly, combine all
- Then match this phase with the input phase – check probabitility and assign name according to it

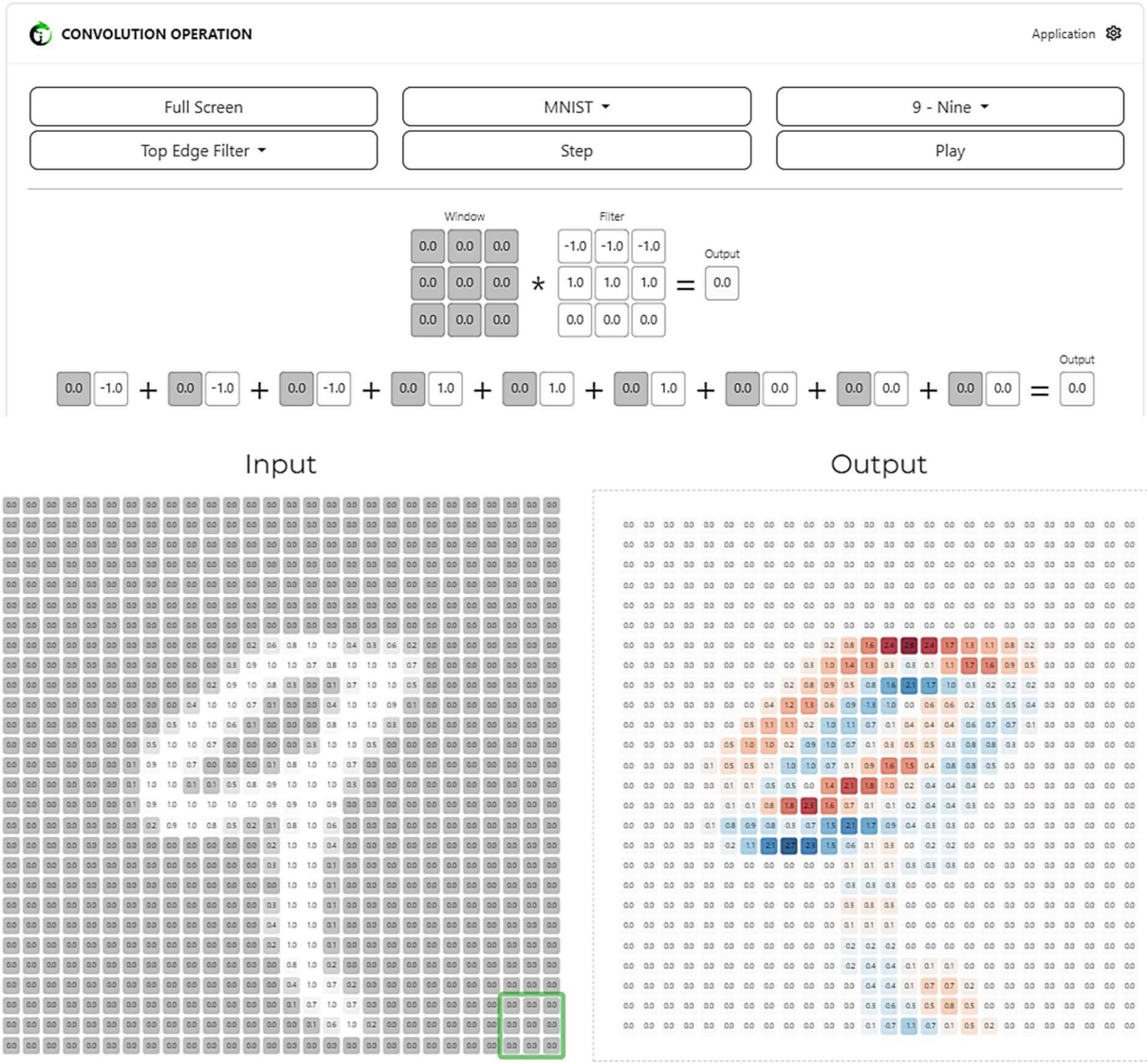
Cosine similarity in vector embedding

Cosine-similarity is the cosine of the angle between two vectors, or equivalently the dot product between their normalizations.

Cosine similarity measures the similarity between two vectors of an inner product space. It is measured by the cosine of the angle between two vectors and determines whether two vectors are pointing in roughly the same direction.



MNIST Number Image at: https://deeplizard.com/resource/pavq7noze2



Filter is 3 x3 (Filter can be vertical and horizontal)

Application of CNN

Image Classification

Mite progress bar is greater as compare to others

Terms

Image Classification

Definition: Image classification is the process of assigning a label or category to an entire image based on its content.

Example: Given an image of a cat, a classification model might output "cat" as the label. It doesn't identify where the cat is within the image; it simply determines that the image contains a cat.

Use Case: This is often used in applications like sorting photos into albums, detecting objects in images for cataloging, or categorizing images for search engines.

Image Identification

Definition: Image identification is a more specific type of classification that often involves recognizing individual objects or entities within an image, sometimes with unique identifiers.

Example: In security systems, image identification might recognize a specific person's face from a database of known faces, rather than just classifying them as a "person.

Use Case: It's used in applications such as personal identification systems, license plate recognition, and branded product detection.

Image Segmentation

Definition: Image segmentation involves dividing an image into multiple segments or regions, often to identify objects, boundaries, or other relevant parts of the image.

Example: Segmenting an image of a street into separate regions like road, cars, pedestrians, and buildings. This allows for more detailed analysis of each part of the image.

Use Case: Useful in medical imaging to highlight specific organs or tissues, in autonomous vehicles to detect and understand road features, and in agriculture to monitor crop health.

Facial Recognition

Definition: Facial recognition is a specific application of image identification focused on recognizing and verifying individuals based on their facial features.

Example: Facial recognition systems can identify a person in a crowd by comparing their face against a database of known faces.

Use Case: Used in security for unlocking devices, accessing secure areas, and surveillance systems, as well as in social media for tagging people in photos.

Summary

- **Image Classification:** Assigns a label to an entire image.
- **Image Identification:** Identifies specific objects or entities, often with unique identifiers.
- **Image Segmentation:** Divides an image into segments or regions for detailed analysis.
- **Facial Recognition:** Identifies or verifies individuals based on facial features.

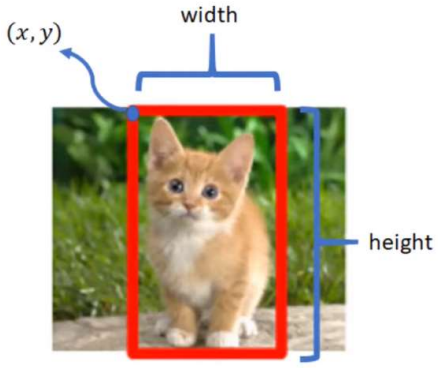
Each technique serves a distinct purpose and is chosen based on the specific needs of the application.

Image Localization

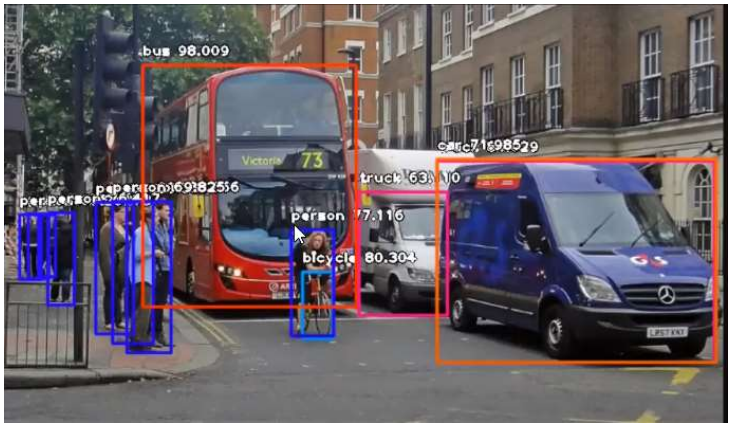
Image localization is a computer vision task focused on identifying and specifying the location of objects or regions of interest within an image. It provides more detailed information than simple image classification by not only recognizing what is present in an image but also pinpointing where it is.

Key Concepts of Image Localization

- ***Bounding Boxes:*** The most common method of localization involves drawing bounding boxes around objects of interest. A bounding box is a rectangular frame that encloses the object.
- ***Coordinate Coordinates***: Localization often involves outputting coordinates that define the location of the object. These can be in the form of bounding box coordinates.
- ***Region Proposal Networks (RPN):*** In more advanced systems, region proposal networks are used to suggest possible regions where objects might be located.



Object Detection:



Note: It tell percentage match from original data

Object detection is a computer vision task that involves not only recognizing objects within an image but also determining their precise locations.

This task combines elements of both image classification and localization to identify and localize multiple objects within a single image.

Key Concepts of Object Detection

- ***Bounding Boxes:*** Object detection often involves drawing bounding boxes around each detected object. These boxes are defined by coordinates (usually the top-left corner and the bottom-right corner) that specify the object's location in the image.
- ***Class Labels:*** Each detected object is assigned a class label that indicates what the object is. For example, if the object is a car, the label would be "car."
- ***Confidence Scores:*** Object detection models typically output confidence scores that indicate the likelihood that the detected object belongs to a certain class. This helps in filtering out less certain detections.
- ***Multiple Objects:*** A robust object detection system can identify and localize multiple objects within a single image, handling overlapping objects and varying object sizes.

Facial Recognition

With the help of point it make grid, and map image on grid according to threshold percentage.

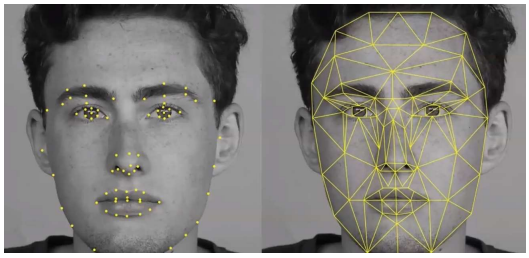
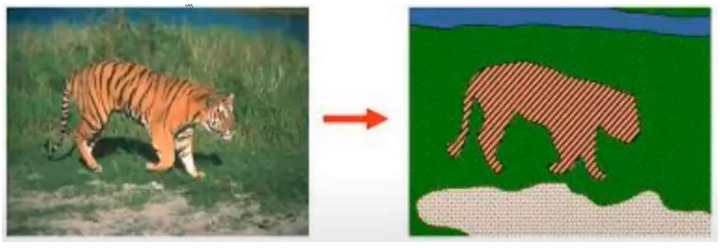


Image Segmentation

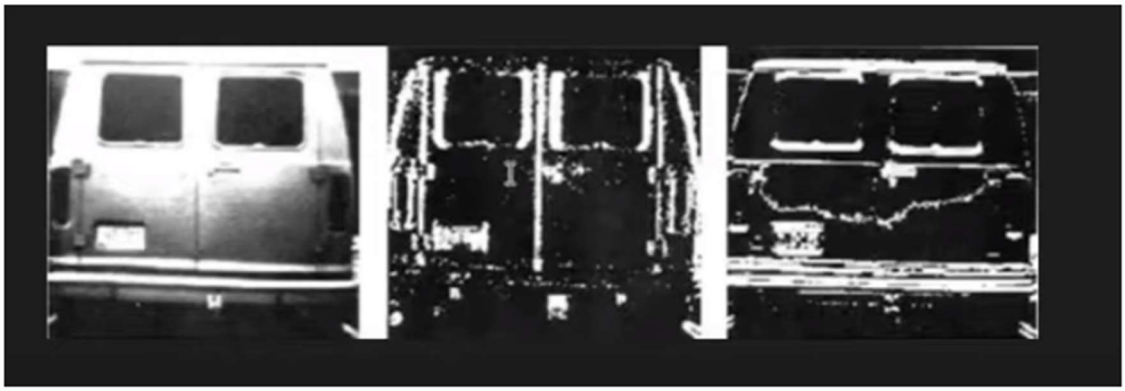


Pose Estimation

Pose estimation is a computer vision task that involves determining the orientation and position of an object or a human body in an image or video. It's used to understand the spatial arrangement and movement of objects or body parts.

Pose estimation can be broadly categorized into **object pose estimation** and **human pose estimation.**

Convolution Operation (Edge Detection)



Vertical(Edge Detection) Horizontal (Edge Detection)

Edge detection is of two types: Horizontal edge detection and Vertical edge detection

Applying Horizontal Edge Detection Kernel

When you apply the horizontal edge detection kernel, you'll notice that the edges in the horizontal direction (i.e., where there is a transition from 1 to 0) are highlighted.

Applying Vertical Edge Detection Kernel

When you apply the vertical edge detection kernel, you'll notice that the edges in the vertical direction (i.e., where there is a transition from 1 to 0) are highlighted.

$$\text{Horizontal Edge Detection} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

$$\text{Vertical Edge Detection} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

Example

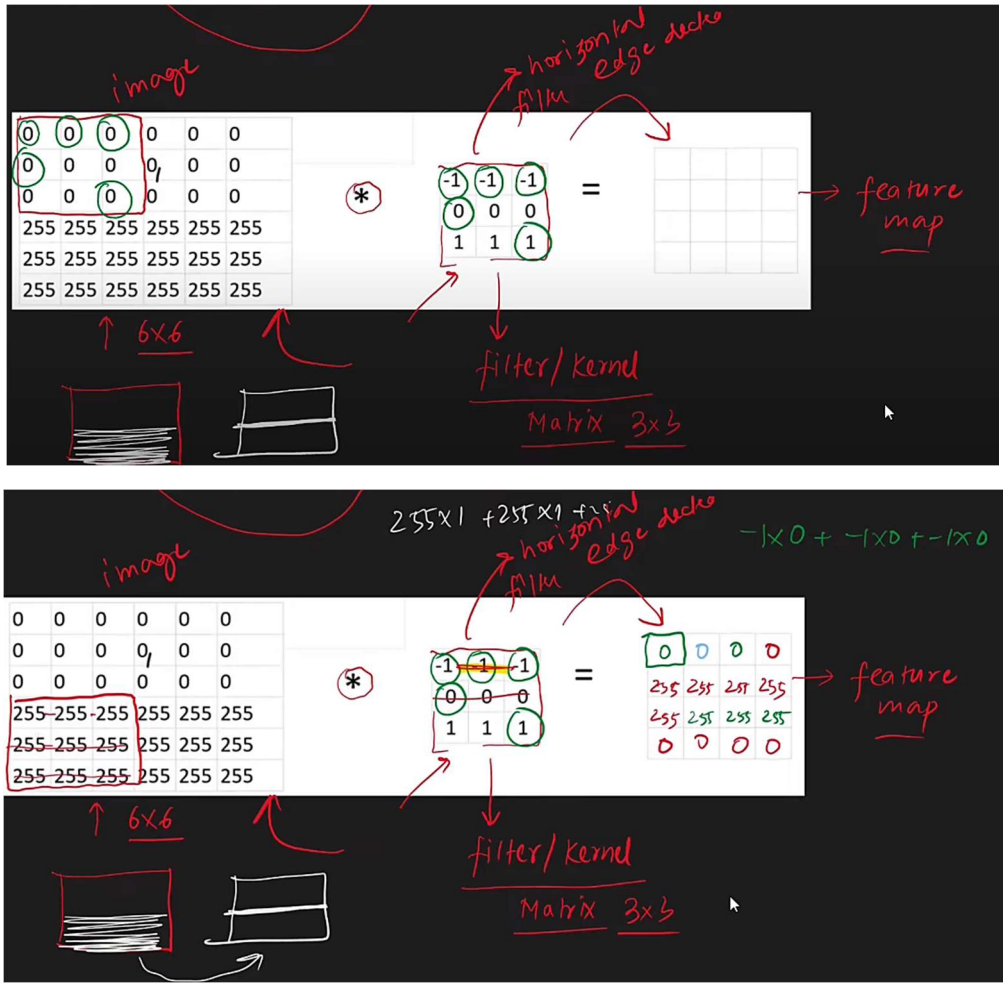
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
255	255	255	255	255	255
255	255	255	255	255	255
255	255	255	255	255	255

*

-1	-1	-1
0	0	0
1	1	1

=

Filter size 3x3 and Image Size 6x6



First image 6x6 – then filter 3x3 – then output

- Filter (Gaussian filter) add on image – move step by step horizontally
- White region appear due to 255 answers came due to filter
- **Problem:** Filter see center information efficiently but see edges only one time – means it is biased – to resolve this introduce padding and stride

Filter in CNN

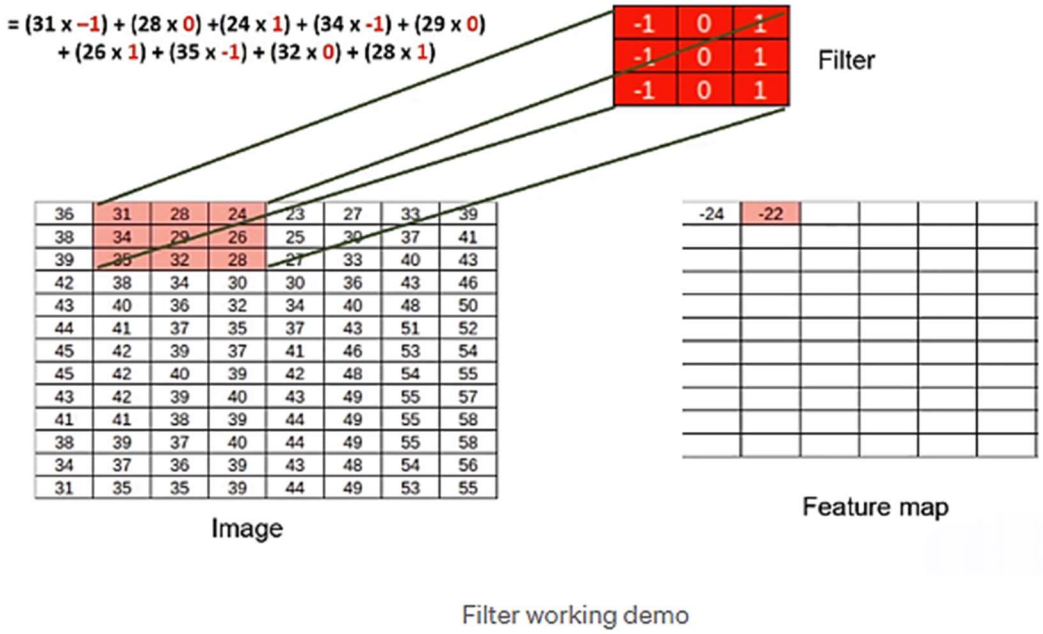


Image shows Vertical Filter

When filter is on the first patch CNN performs element wise multiplication of source image value and filter and then sums it up. This value becomes the value of feature map. Once the CNN is done with calculating the value for the particular patch, the filter moves to the next patch horizontally and repeats this process till the end of the image.

Key filters used in Image processing

Gaussian Filter: Applies a Gaussian function to smooth an image or signal, reducing noise and detail.

Edge Detection Filter: Identifies and highlights the edges in an image by detecting rapid changes in intensity.

Vertical Filter: Emphasizes or processes features oriented vertically in an image or signal.

Horizontal Filter: Emphasizes or processes features oriented horizontally in an image or signal.

Angle Filter: Targets features or components of a signal or image based on a specific angle or orientation.