Convolutional Neural Network (CNN)

CNN → Images, videos → visual data

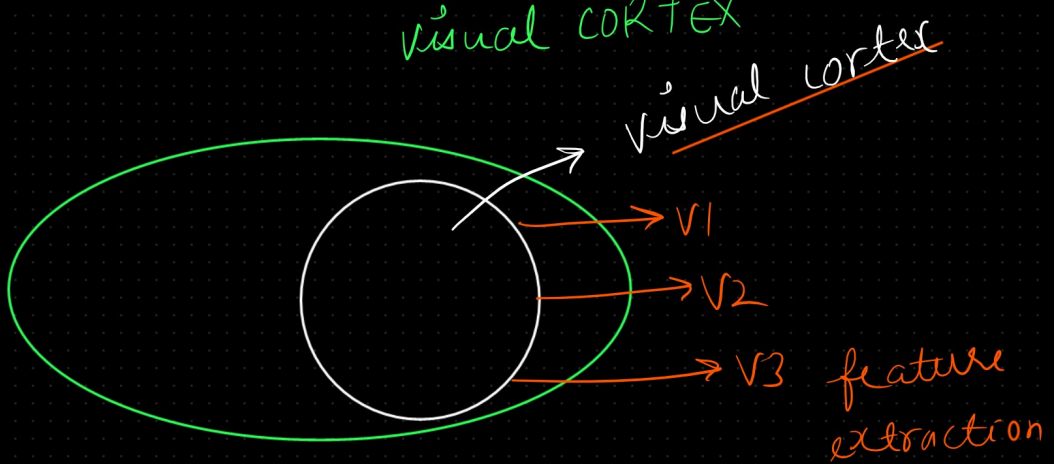
A Convolutional Neural Network (CNN) is a type of deep learning neural network that is well-suited for image and video analysis. CNNs use a series of convolution and pooling layers to extract features from images and videos, and then use these features to classify or detect objects or scenes.

→ Facial Recognition
→ Image Analysis
→ Object Detection

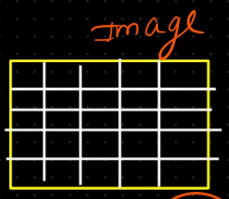
} Computer vision

How does a person recognize different objects

human Brain
↓
visual CORTEX



Images → Matrix → rows and columns



5×5 pixel

Images

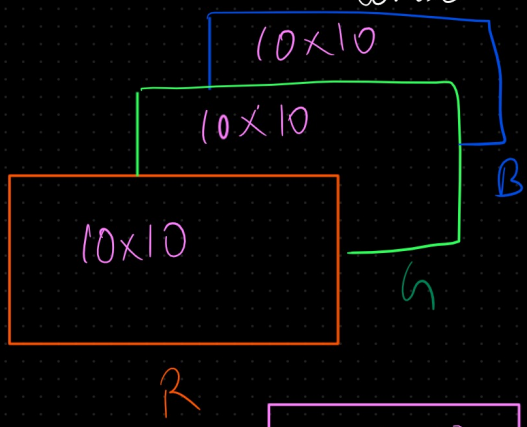
- Grayscale (Black, white) → 1 channels → single channel → 5×5×1 → gray
- Color Image (RGB) → Red, Green, Blue → 3 channels → Multi channel → 5×5×3

0 - 255 → Range

0 ↓ Black — 255 ↓ White

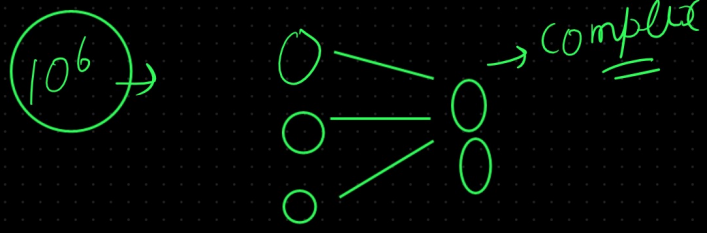
GRAY

5×5×1

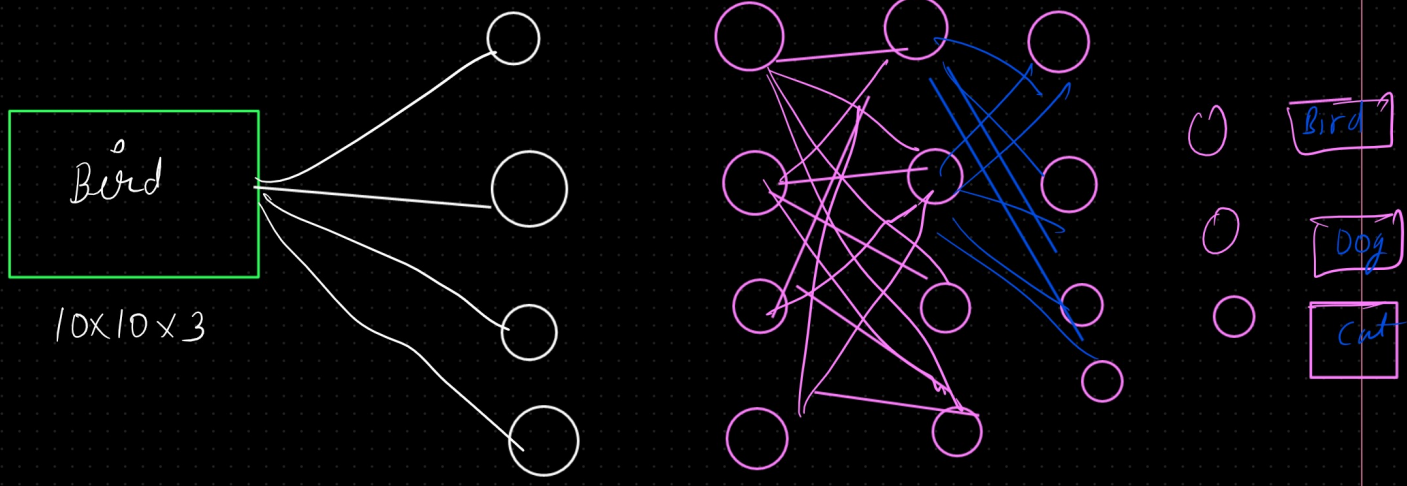


10×10×3

Images → 1000 × 1000 pixels



Extensive use CPU
GPU



Aim: We need to classify the Image

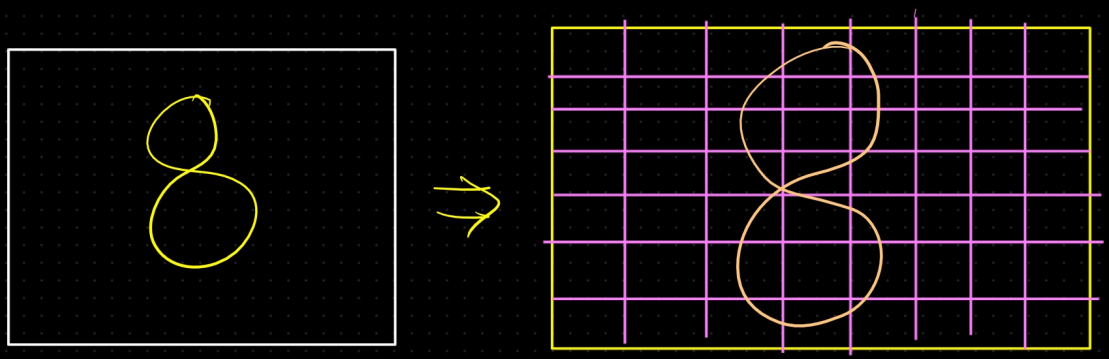
CNN → Multiple Hidden layers

feature extractions → different calculations

- Convolution layer
- Relu layer
- pooling layer

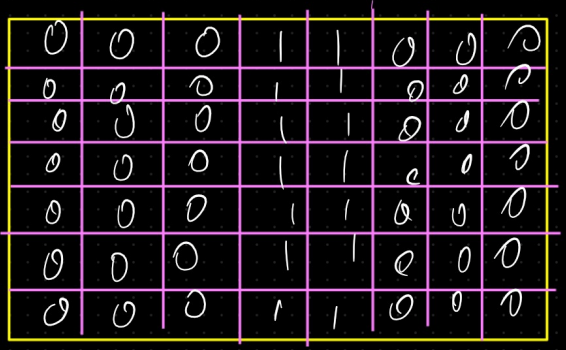
} perform feature extraction

A convolutional neural network is a feed-forward neural network that is generally used to analyze visual images by processing data with grid-like topology. It's also known as a ConvNet. A convolutional neural network is used to detect and classify objects in an image.



Image

Array Resp^{n}



pixels form

Convolutional Neural Network

⇓

convolution operation

$a = [5, 3, 7, 5, 9, 7]$

$b = [1, 2, 3]$

× multiply element wise and then product

$a * b$

$a = [5, 3, 2, 5, 9, 7]$

$b = [1, 2, 3]$

$a * b = 5(1) + 3(2) + 2(3)$

$= 5 + 6 + 6 = 17$

$b = [1, 2, 3]$

$a * b = 3(1) + 2(2) + 5(3)$

$= 22$