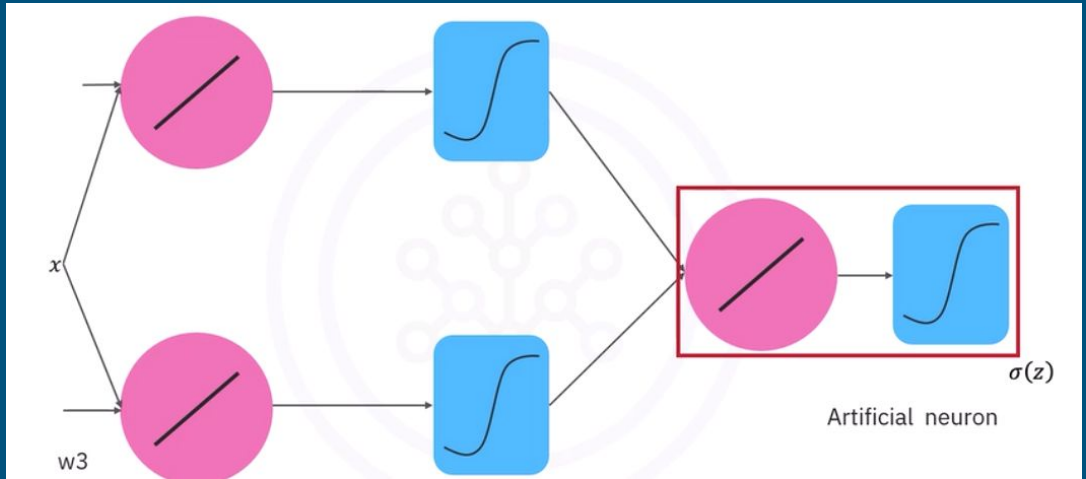
IMAGE PROCESSING

WEEK-1

5. NN to classify image

- **Neural network basics:**

- A neural network is a computational model inspired by human brain.
- It is used for separating non-linearly separable data.
- The work of neural network is to approximate the class by using learnable parameter.
- In this diagram it represents a two-layer neural network, We have the hidden layer.
- Each linear function and activation function is known as an artificial neuron.
- In this figure the hidden layer has two artificial neurons.
- The output layer has one artificial neurons

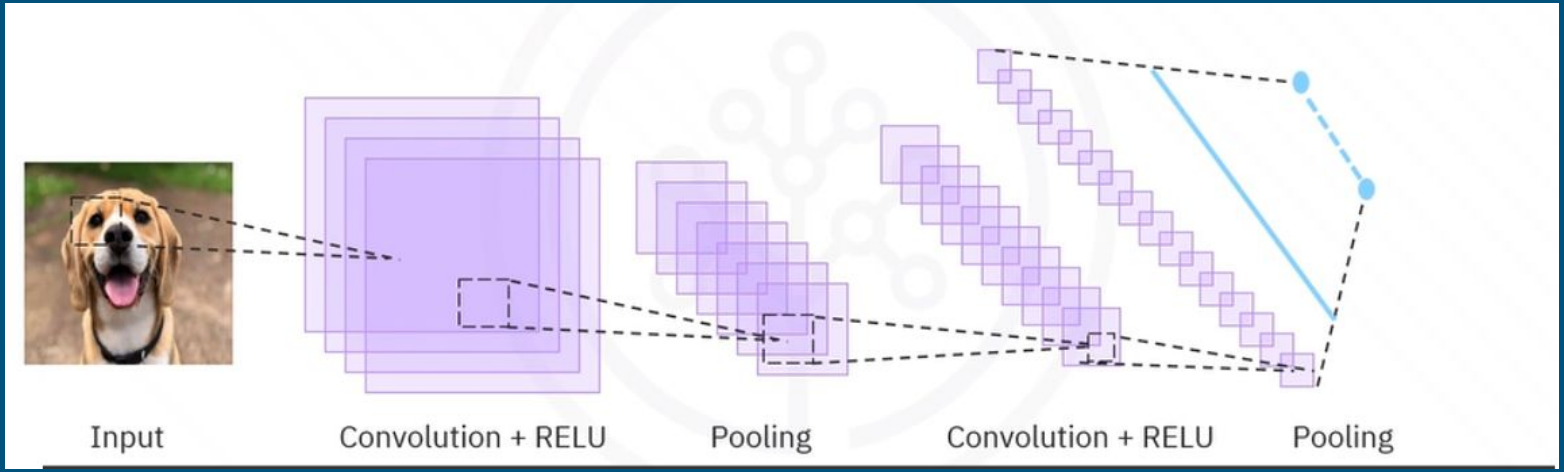


- Fully connected Neural network architecture :

- Fully connected neural network architecture deals with how to arrange the different number of hidden layers and neurons.
- In the neural network if we have more than one hidden layers that is called a deep neural network.
- But neurons, or more layers may lead to overfitting.
- The output layer uses activation function .Each neuron is a linear classifier
- To learn parameters we have to perform gradient descent , but deeper the network gets the smaller the gradient gets.This is called vanishing gradient problem
- Neural network are trained in a similar manner to logistic regression and softmax

- Sigmoid activation function suffers from vanishing gradient problem but ReLU solve the vanishing gradient problem.

- What is CNN?
- A convolutional network or CNN is a neural network with special layers.
- The model classifies an image by taking a part of the image . Each input image will pass through series of convolutional lakers with filter ,pooling layers,fully connected layers and while applying activation functions to classify an object



- Convolution and pooling layers are the first layers used to extract features from input.
- Each kernel or filter of Convolution will detect different property of the image.
- Receptive field id the size of the region in the input that produces a pixel value in av.

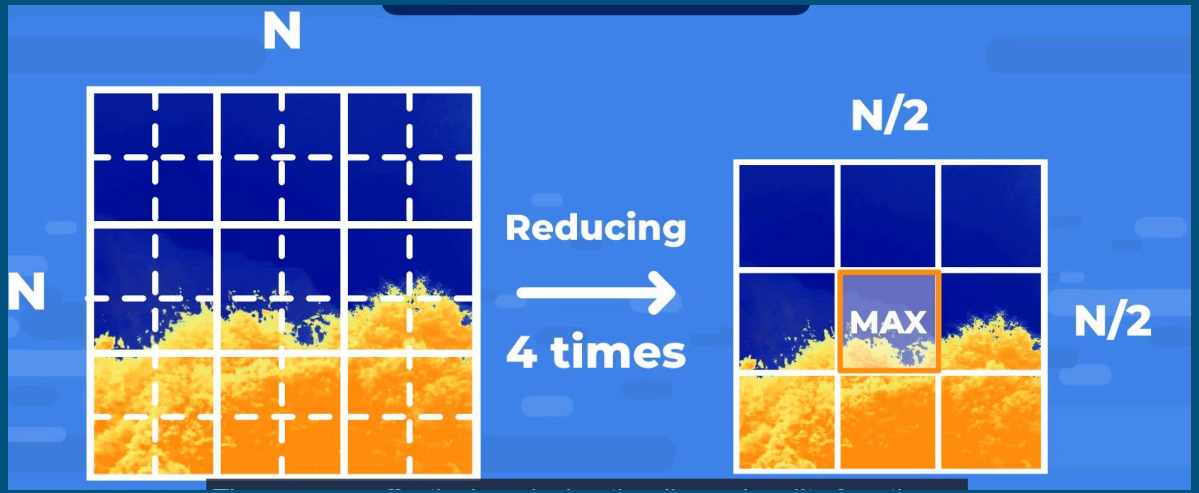
- **Why CNN?**

- CNN's are used as they understand an image just an human brain does.
- Characteristics:
  - **They captures spatial features :** as cnn use convolutions(filters) to capture the local patterns automatically.
  - **Automatic feature extraction :** Traditional ML were required to use : HOG,SIFT,SURF,LBP but CNN learned to extract features automatically, and no manual feature engineering needed.
  - **Parameter sharing :** as a 3X3 filter slides is used across the whole image , this reduces the no of parameters as 3X3 filter only requires 9 parameters . This makes CNN faster, cheaper.
  - **Translation invariance :** if in an image an object is slightly moved , a normal nn will fail but a CNN will still recognizes it as the filters detects features
  - **Work extremely well in real-world vision tasks :** is used for real-world application as it outperforms classical ML by a huge margin
  - **Hierarchy of feature feature learning :** CNN learns and extracts features from simple to complex as hierarchy.
  - **Efficient with images,videos,audio :** CNN works well with grid-like data.

- **Convolutional -> filter :**

- **Filter :** a filter also known as kernel is a small matrix of numbers of dimensions usually : 3X3 , 5X5 , 7X7
- Convolution is the sliding of the filter over the image and multiplying pixel values and output is called feature map.
- **In CNN we use convolution for detecting patterns like:**
  - Edge
  - Corners
  - Texture
  - Color change
  - Curve
  - Shape part
- Different filter learn different features.
- In classical computer vision, human designed filters but in CNN filters automatically learns values

- **Max pooling :**
- Max pooling is an downsampling operation used in CNN.
- **It helps in :**
  - Reduces image size
  - Reduce computation
  - Keep important features
  - Make the network more robust
- Most commonly we use 2X2 max pooling with stride 2
- In max-pooling we select the maximum number from the selected matrix



- **Activation functions :**

- An activation function decides whether a neuron is activated or not.
- It helps in introducing non-linearity.
- Without activation function becomes a neural network becomes just a linear equation
- Common activation functions:
- **ReLU(rectified linear unit):**
  - ReLU is commonly used in modern neural network and especially in CNN
  - Formulae:

$$\text{ReLU}(x) = \max(0, x)$$

  - It works by : if input is positive then it return the input but if its is negative or zero it returns 0.
  - It helps in removing negative values and helps in avoiding vanishing gradient
  - Drawback : dying ReLU problem : sometimes neurons get stuck and always return 0 as output
  - Relu is not used at the output layer.

- Softmax:

- Softmax function is an activation function used in the output layer of multi-class classification problems.
- Network normally outputs number and not probllaties , to arrange the output in the range of 0 - 1 such that sum = 1
- Formulae :

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}}$$

- Input is converted into probabilities .
- Used in Multi-class classification.

• Architecture Overview(VGG,ResNet) :

- VGG architecture(VGG16/VGG19) :

- Stands for visual geometry group created by oxford VGG in 2014.
- To shows the deep networks with small filters (3X3) improves accuracy.
- **Key components :**
- A. uses only 3X3 convolutions : kernel size->3x3 , stride = 1,easy to stack
- B. Uses max-pooling : reduces spatial size and keeps depth
- C. Depth : VGG16 ->16 layers , VGG19 -> 19 layers
- **Drawbacks :** very heavy(138 parameters) , slow to train , large memory usage

- ResNet :

- Stands for Residual networks introduced by microsoft in 2015.
- It is used for solving the vanishing gradient problem when CNN becomes very deep
- This network skips a layer if not needed
- Helps in preventing vanishing gradients , Degradation in accuracy and hard optimization.

ResNet learns:

$$F(x) = H(x) - x$$

And output:

$$H(x) = F(x) + x$$

- POC:
- Training miniature CNN model for cats and dogs classification
- https://colab.research.google.com/drive/10dxQFzejMrkO_1PspD09JPS959bQMmLs#scrollTo=K0qeIqFe57wH