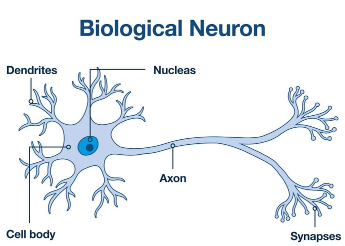
Deep Learning Walk Through

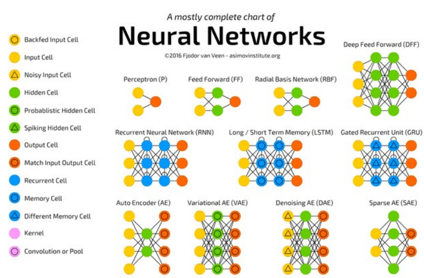
Dr. Shailesh Sivan
Assistant Professor
DCS, CUSAT

PERCEPTRON

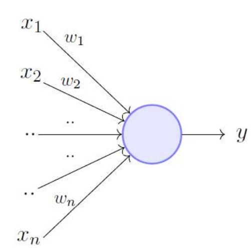
- First neural network learning model in the 1960's
- Simple and limited (single layer models)
- Basic concepts are similar for multi-layer models so this is a good learning tool
- Still used in many current applications (modems, etc.)

NEURAL NETWORKS





PERCEPTRON MODEL



$$y = 1 \quad if \sum_{i=1}^{n} w_i * x_i \geq \theta$$

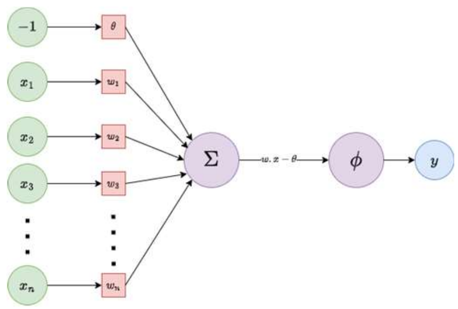
$$= 0 \quad if \sum_{i=1}^{n} w_i * x_i < \theta$$

Rewriting the above,

$$y = 1 \quad if \sum_{i=1}^{n} w_i * x_i - \theta \geq 0$$

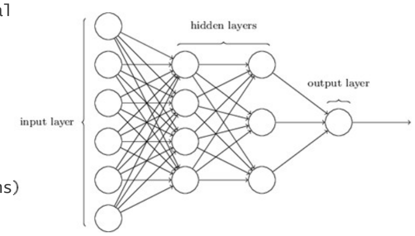
$$= 0 \quad if \sum_{i=1}^{n} w_i * x_i - \theta < 0$$

PERCEPTRON MODEL

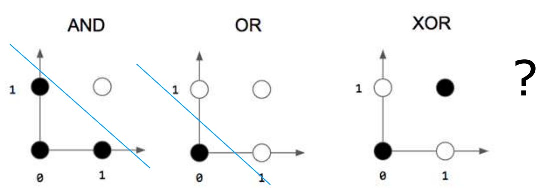


MULTI LAYER PERCEPTRON(MLP)

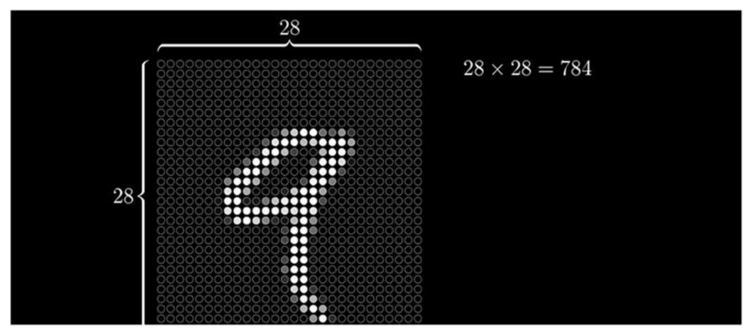
- Feedforward network: The neurons in each layer feed their output forward to the next layer until we get the final output from the neural network.
- There can be any number of hidden layers within a feedforward network.
- The number of neurons can be completely arbitrary.
- MLP used to describe any general feedforward (no recurrent connections) network



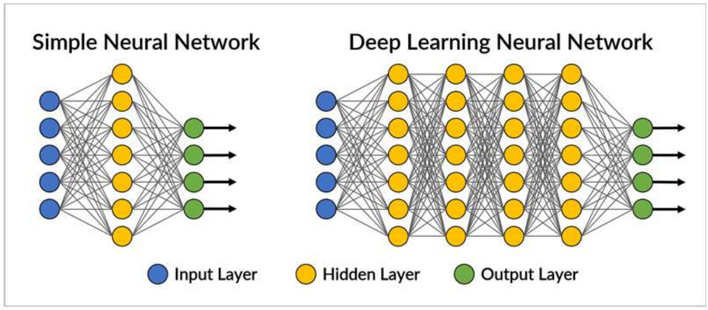
MORE GATES



MULTI LAYER PERCEPTRON(MLP)

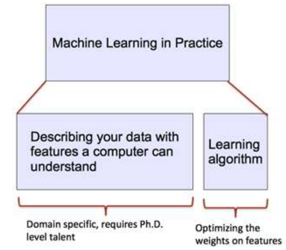


DEEP NEURAL NETWORK

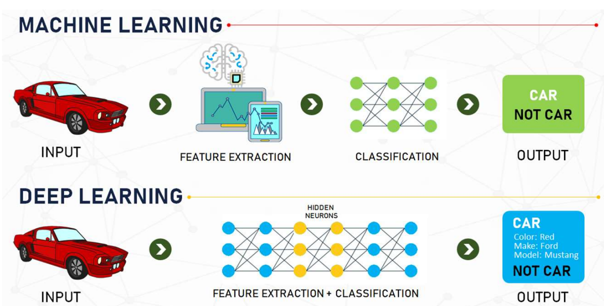


MACHINE VS DEEP LEARNING

- Most machine learning methods work well because of human-designed representations and input features
- ML becomes just optimizing weights to best make a final prediction



DEEP LEARNING



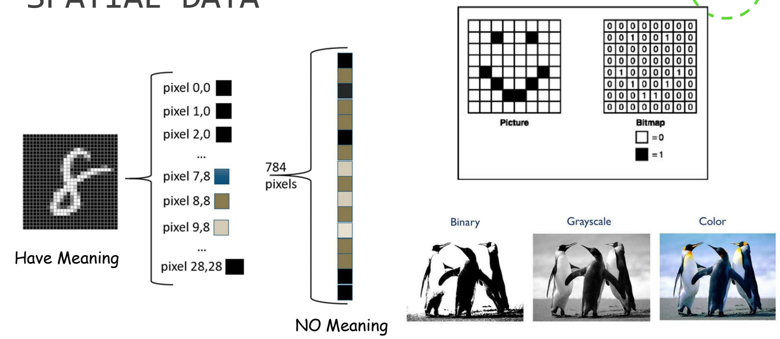
MACHINE VS DEEP LEARNING

- A machine learning subfield of learning representations of data. Exceptional effective at learning patterns.
- Deep learning algorithms attempt to learn (multiple levels of) representation by using a hierarchy of multiple layers
- If you provide the system tons of information, it begins to understand it and respond in useful ways.

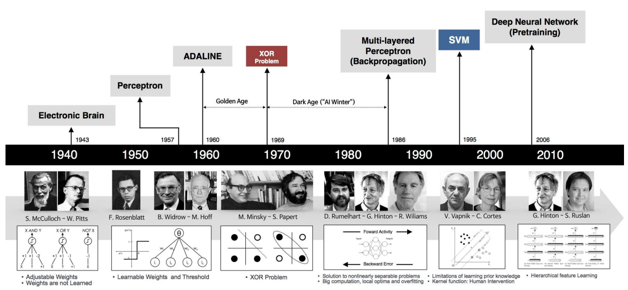
WHY DEEP LEARNING USEFUL?

- Manually designed features are often over-specified, incomplete and take a long time to design and validate
- Learned Features are easy to adapt, fast to learn
- Deep learning provides a very flexible, (almost?) universal, learnable framework for representing world, visual and linguistic information.
- Can learn both unsupervised and supervised
- Effective end-to-end joint system learning
- Utilize large amounts of training data

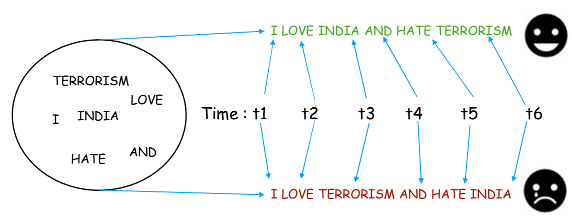
SPATIAL DATA



HISTORY OF DEEP LEARNING



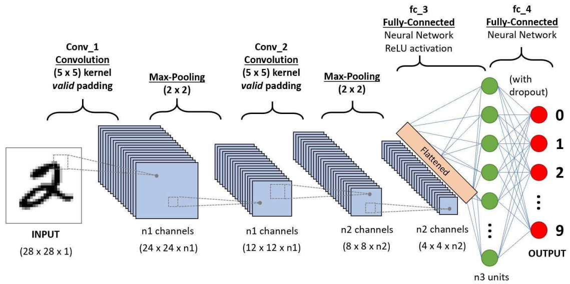
TEMPORAL DATA



"There is nothing to like in this movie."

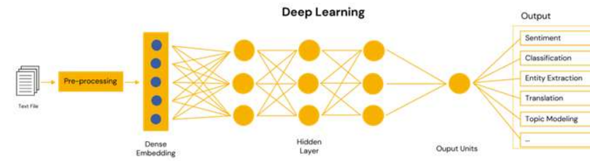
AGCCCCTGTGAGGAACTAG

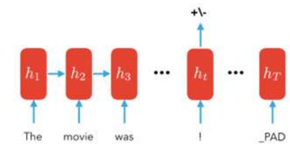
DEEP LEARNING(VISION)



Convolutional Neural Networks

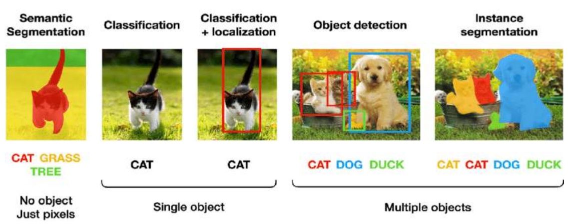
DEEP LEARNING(NLP)



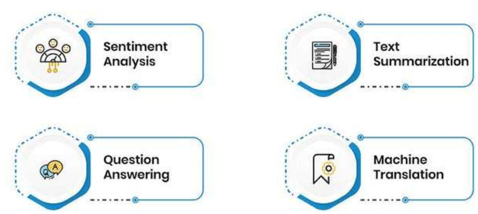


Reccurent Neural Networks

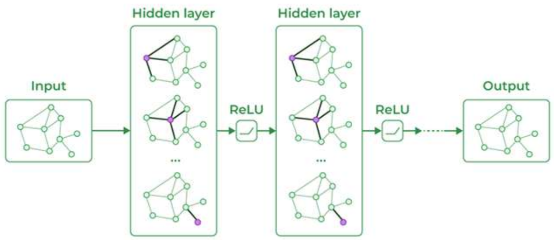
DEEP LEARNING FOR VISION



DEEP LEARNING FOR VISION

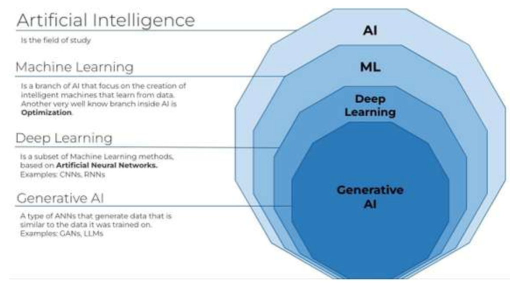


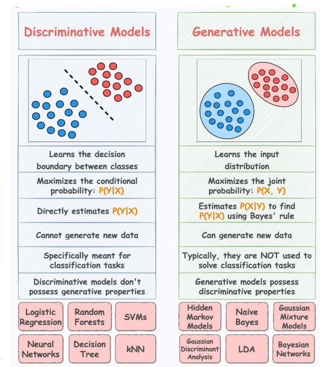
DEEP LEARNING(GRAPH)



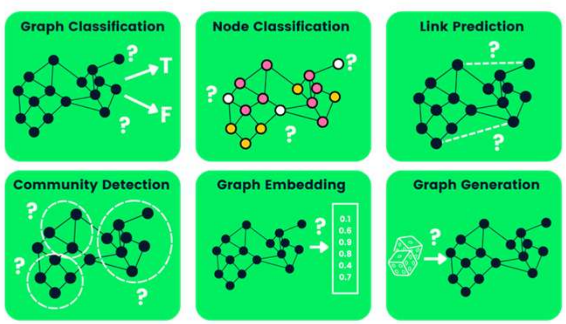
Graph Neural Networks

GENERATIVE AI/MODELS

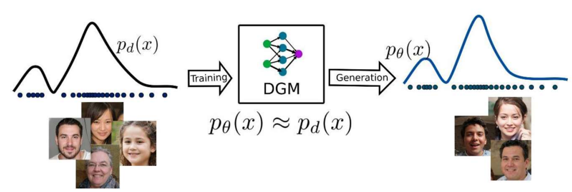




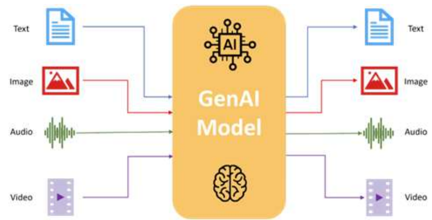
DEEP LEARNING FOR GRAPHS



Deep Generative Models

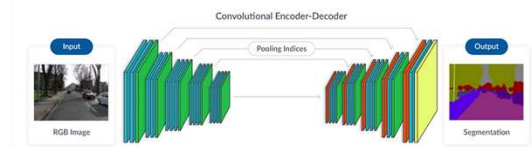


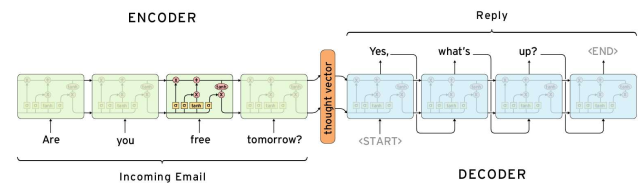
GENAI MODELS : AN OVERVIEW



- Encoder – Decoder Models
- Generative Adverserial Networks
- Transformers
- Vision Transformers
- Diffusion Models

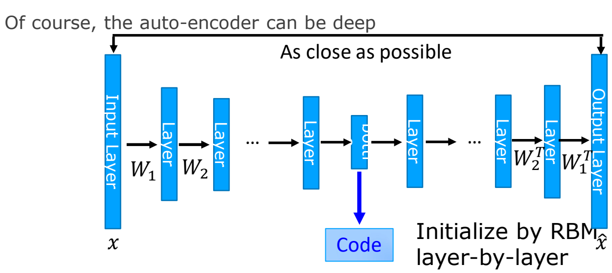
ENCODER-DECODER



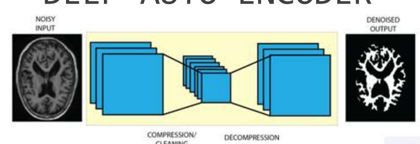


DEEP AUTO-ENCODER

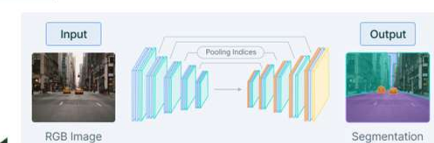
Symmetric is
not necessary.



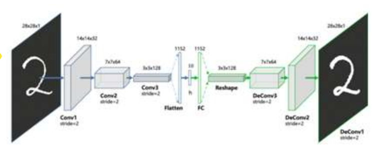
DEEP AUTO-ENCODER



Denoisin
g

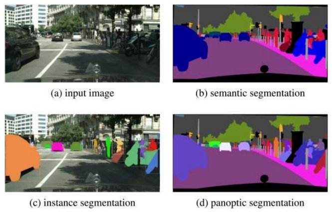


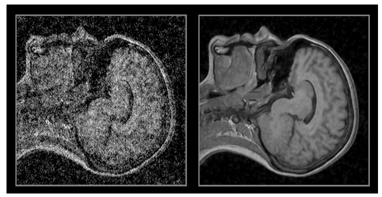
Segmentati
on

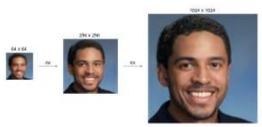


Associati
ve

ENCODER-DECODER(VISION)

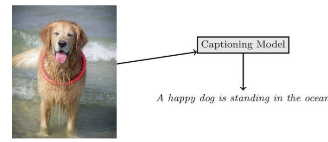


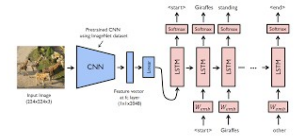


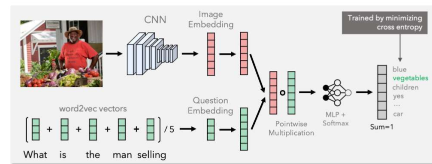


ENCODER-DECODER(TEXT-VISION)

IMAGE CAPTIONING

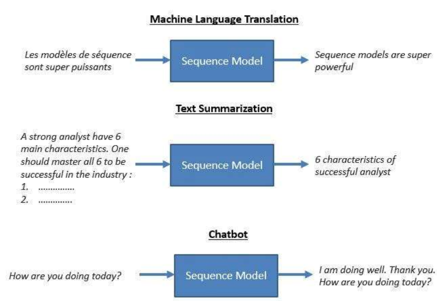




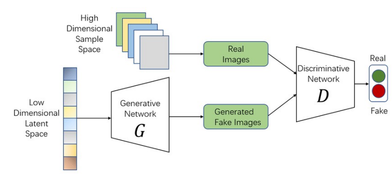


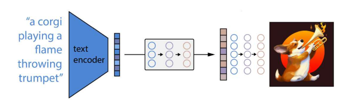
VISUAL QUESTION ANSWERING

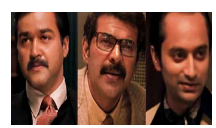
ENCODER-DECODER(TEXT)

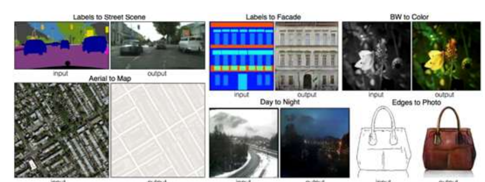


GENERATIVE ADVESERIAL NETWORKS(GANS)

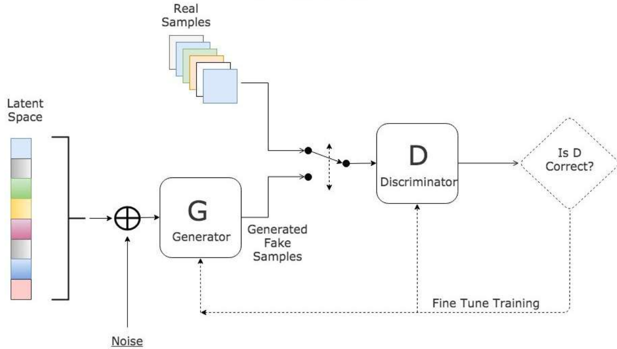




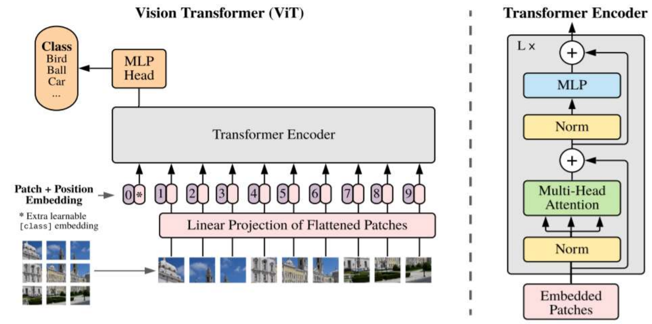




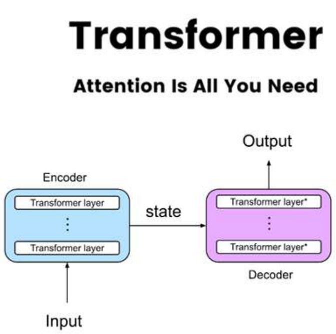
GENERATIVE ADVERSERIAL NETWORK

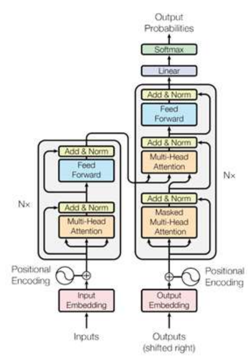


THE SECRET : TRANSFORMER



THE SECRET : TRANSFORMER





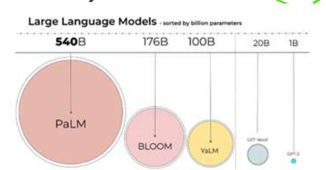
TRANSFORMER TO GPT(AN LLM)

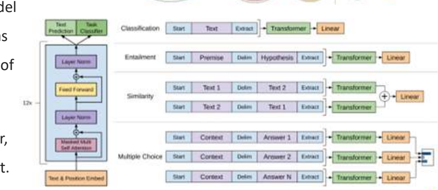
Generative Pretrained Model(GPT)

GPT-3 (Generative Pre-trained Transformer 3) is a language model that was created by **OpenAI**.

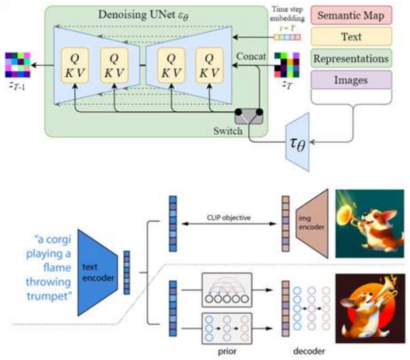
The 175-billion parameter deep learning model capable of producing human-like text and was trained on large text datasets with hundreds of billions of words.

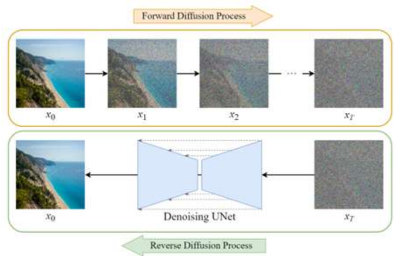
GPT uses an unmodified Transformer decoder, except that it lacks the encoder attention part.





DIFFUSION

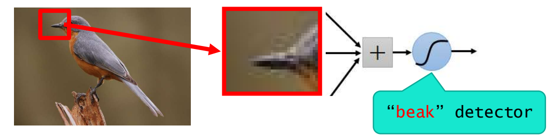




CONSIDER LEARNING AN IMAGE:

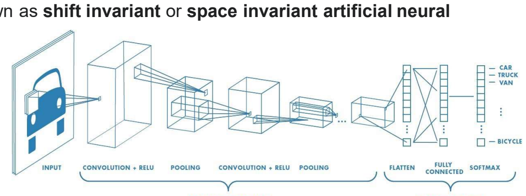
Some patterns are much smaller than the whole image

Can represent a small region with fewer parameters



CONVOLUTIONAL NEURAL NETWORKS

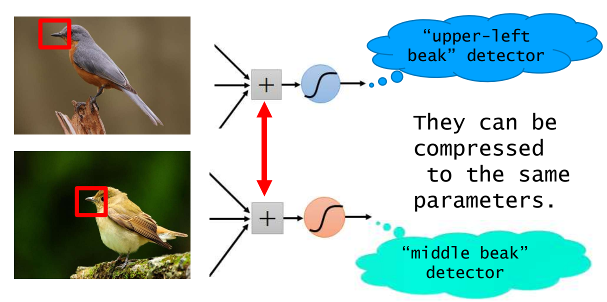
- A convolutional neural network (CNN, or ConvNet) is a class of artificial neural network, most commonly applied to analyze visual imagery
- They are also known as **shift invariant** or **space invariant artificial neural networks** (SIANN)



38

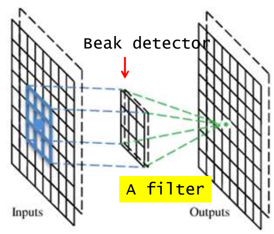
SAME PATTERN APPEARS IN DIFFERENT PLACES: THEY CAN BE COMPRESSED!

WHAT ABOUT TRAINING A LOT OF SUCH "SMALL" DETECTORS AND EACH DETECTOR MUST "MOVE AROUND".



A CONVOLUTIONAL LAYER

A CNN is a neural network with some convolutional layers (and some other layers). A convolutional layer has a number of filters that does convolutional operation.



CONVOLUTION

stride=1

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image

Dot product → 3 -1

Filter 1

1	-1	-1
-1	1	-1
-1	-1	1

CONVOLUTION

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image

These are the network parameters to be learned.

-1	1	-1
-1	1	-1
-1	1	-1

Filter 1

1	-1	-1
-1	1	-1
-1	-1	1

Filter 2

⋮

Each filter detects a small pattern (3 x 3).

CONVOLUTION

If stride=2

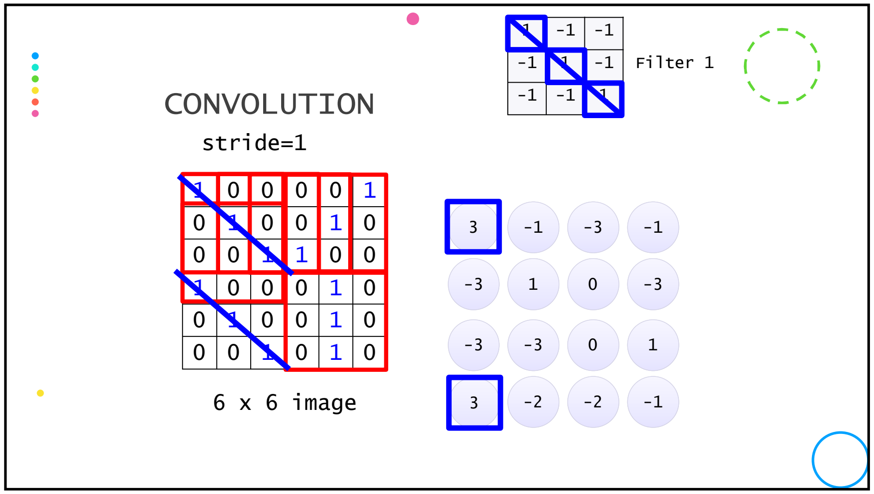
1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

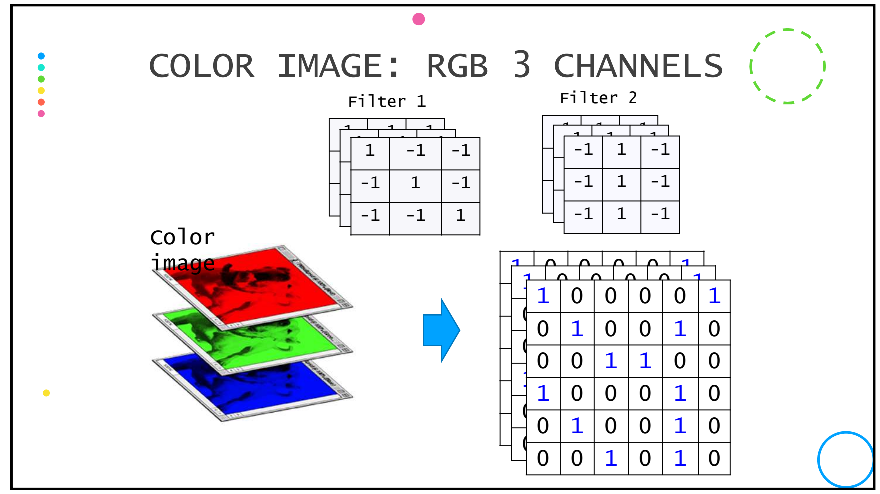
6 x 6 image

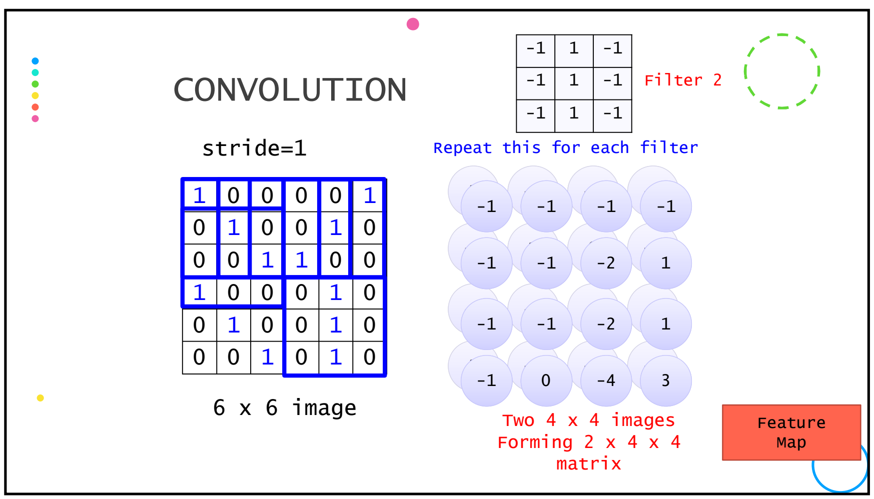
3 -3

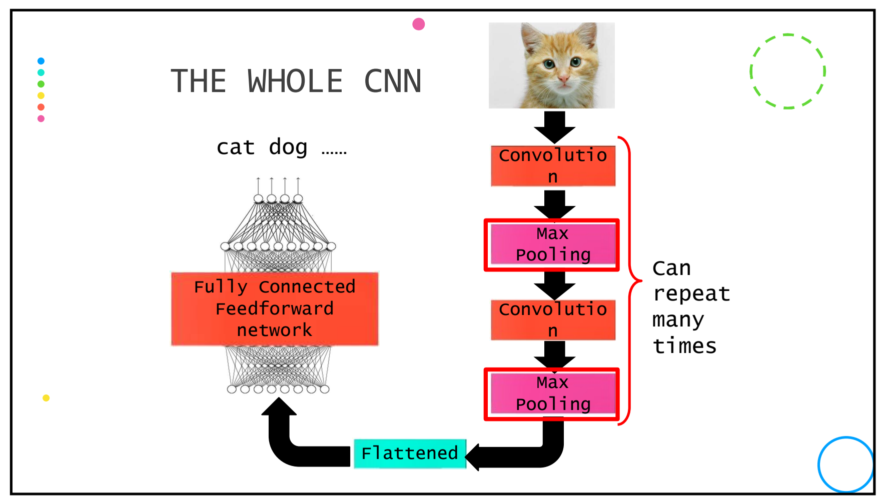
1	-1	-1
-1	1	-1
-1	-1	1

Filter 1


CONVOLUTION
stride=1
Filter 1
1 -1 -1
-1 1 -1
-1 -1 1
1 0 0 0 0 1
0 1 0 0 1 0
0 0 1 1 0 0
1 0 0 0 1 0
0 1 0 0 1 0
0 0 1 0 1 0
6 x 6 image
3 -1 -3 -1
-3 1 0 -3
-3 -3 0 1
3 -2 -2 -1


COLOR IMAGE: RGB 3 CHANNELS
Filter 1
1 -1 -1
-1 1 -1
-1 -1 1
Filter 2
-1 1 -1
-1 1 -1
-1 1 -1
Color image
1 0 0 0 0 1
0 1 0 0 1 0
0 0 1 1 0 0
1 0 0 0 1 0
0 1 0 0 1 0
0 0 1 0 1 0


CONVOLUTION
stride=1
-1 1 -1
-1 1 -1
-1 1 -1
Filter 2
Repeat this for each filter
1 0 0 0 0 1
0 1 0 0 1 0
0 0 1 1 0 0
1 0 0 0 1 0
0 1 0 0 1 0
0 0 1 0 1 0
6 x 6 image
-1 -1 -1 -1
-1 -1 -2 1
-1 -1 -2 1
-1 0 -4 3
Two 4 x 4 images
Forming 2 x 4 x 4
matrix
Feature
Map


THE WHOLE CNN
cat dog ……
Fully Connected
Feedforward
network
Convolutio
n
Max
Pooling
Convolutio
n
Max
Pooling
Can
repeat
many
times
Flattened

MAX POOLING

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

-1	1	-1
-1	1	-1
-1	1	-1

Filter 2

3	-1	-3	-1
-3	1	0	-3
-3	-3	0	1
3	-2	-2	-1

-1	-1	-1	-1
-1	-1	-2	1
-1	-1	-2	1
-1	0	-4	3

A CNN COMPRESSES A FULLY CONNECTED NETWORK IN TWO WAYS:

Reducing number of connections

Shared weights on the edges

Max pooling further reduces the complexity

WHY POOLING

Subsampling pixels will not change the object

bird

Subsampling

bird

We can subsample the pixels to make image smaller

fewer parameters to characterize the image

MAX POOLING

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image

Conv

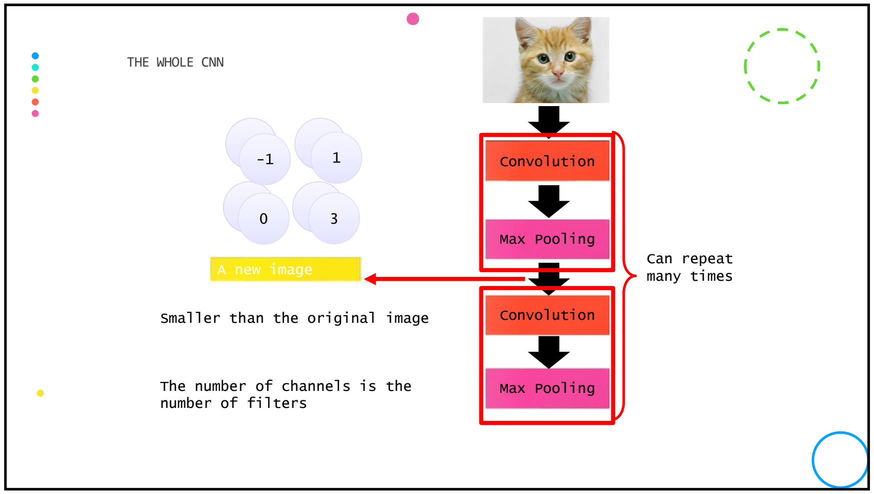
Max Pooling

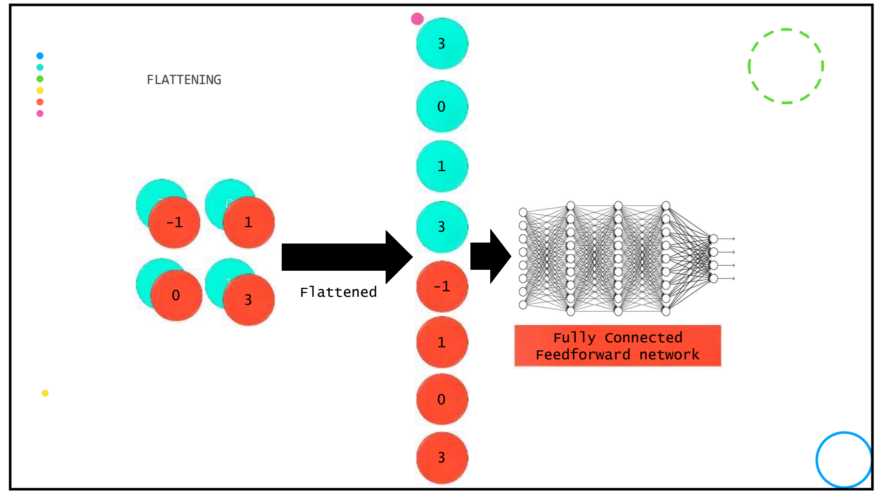
New image but smaller

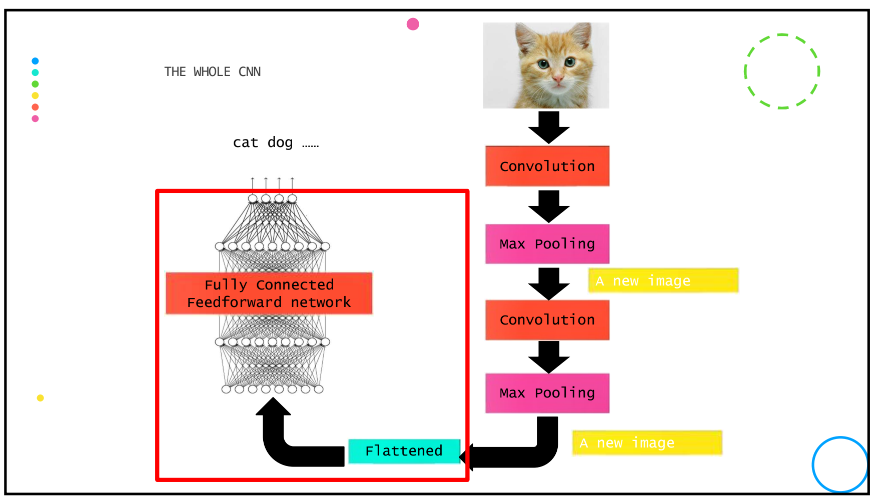
-1	1
0	3

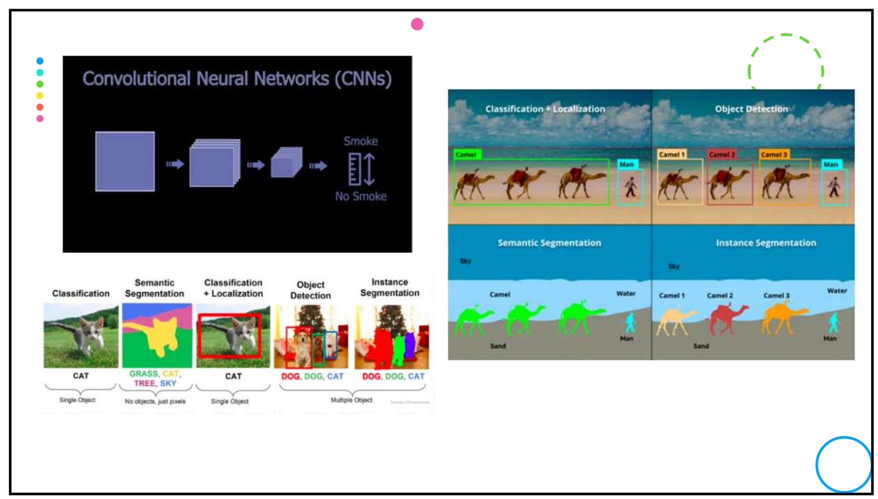
2 x 2 image

Each filter is a channel

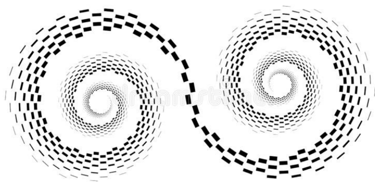

THE WHOLE CNN
-1
1
0
3
A new image
Smaller than the original image
The number of channels is the number of filters
Convolution
Max Pooling
Convolution
Max Pooling
Can repeat many times


FLATTENING
-1
1
0
3
Flattened
3
0
1
3
-1
1
0
3
Fully Connected Feedforward network

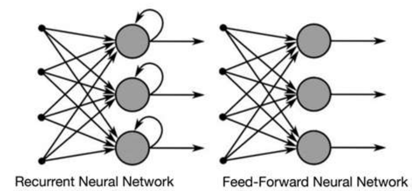

THE WHOLE CNN
cat dog ……
Fully Connected Feedforward network
Flattened
Convolution
Max Pooling
A new image
Convolution
Max Pooling
A new image


Convolutional Neural Networks (CNNs)
Smoke
No Smoke
Classification
Semantic Segmentation
Classification + Localization
Object Detection
Instance Segmentation
CAT
GRASS, CAT, TREE, SKY
CAT
DOG, DOG, CAT
DOG, DOG, CAT
Single Object
No objects, just pixels
Single Object
Multiple Object
Classification + Localization
Object Detection
Semantic Segmentation
Instance Segmentation

RECURRENT NEURAL NETWORKS



Recurrent Nerural Networks



RNN the information cycles through a loop. When it makes a decision, it considers the current input and also what it has learned from the inputs it received previously.

A recurrent neural network, however, is able to remember those characters because of its internal memory. It produces output, copies that output and loops it back into the network.

Simply put: recurrent neural networks add the immediate past to the present.

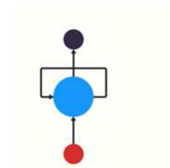
Recurrent Nerural Networks

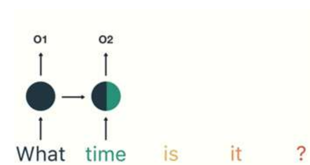
WHAT IS A RECURRENT NEURAL NETWORK (RNN)?

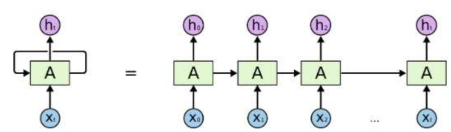
Recurrent neural networks (RNN) are a class of neural networks that are helpful in modeling sequence data. Derived from feedforward networks, RNNs exhibit similar behavior to how human brains function.

Simply put: recurrent neural networks produce predictive results in sequential data that other algorithms can't.

Recurrent Nerural Networks





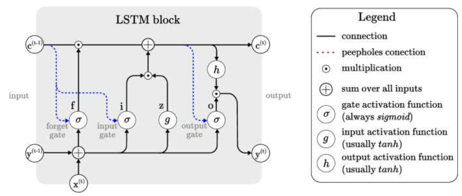


RNN has two inputs: the present and the recent past. This is important because the sequence of data contains crucial information about what is coming next, which is why a RNN can do things other algorithms can't.

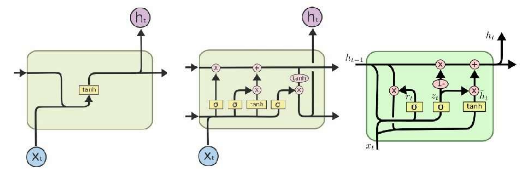
Long Short Term Memory

Long short-term memory networks are an extension for recurrent neural networks, which basically extends the memory. Therefore it is well suited to learn from important experiences that have very long time lags in between.

The units of an LSTM are used as building units for the layers of a RNN, often called an LSTM network



All Together



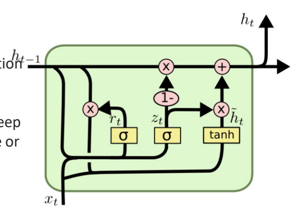
(a) RNN (b) LSTM (c) GRU

Gated Recurrent Units

To solve the vanishing gradient problem of a standard RNN, GRU uses, so-called, **update gate and reset gate**.

Basically, these are two vectors which decide what information should be passed to the output. T

he special thing about them is that they can be trained to keep information from long ago, without washing it through time or remove information which is irrelevant to the prediction.



THANK YOU

ANY QUESTION

Dr. Shailesh Sivan

+91 8907230664

shaileshsivan@cusat.ac.in

https://shaileshsivan.info