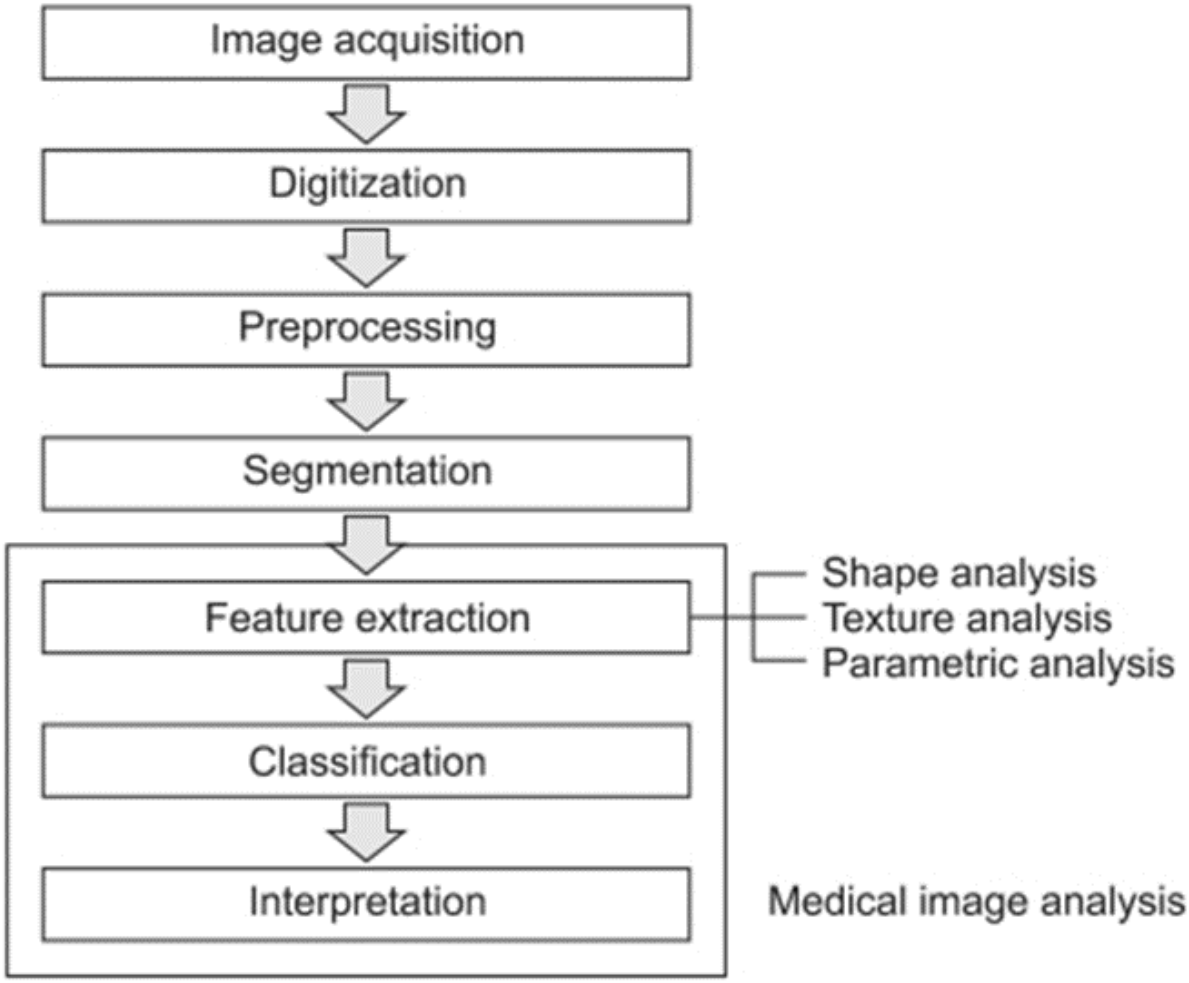
UNIT 1

What is Biomedical Image Processing ?Explain with a block diagram

Biomedical image processing is a technique of extracting required data from the captured images in the medical field.

It includes the analysis, enhancement and display of images captured via x-ray, ultrasound, MRI, nuclear medicine and optical imaging technologies.

Image reconstruction and modelling techniques allow instant processing of 2D signals to create 3D images.



1. Image Acquisition

This is the first step or process of the fundamental steps of digital image processing. Image acquisition could be as simple as being given an image that is already in digital form. Generally, the image acquisition stage involves per-processing, such as scaling, etc.

2. Digitization

Digitization is **the process of converting information into a digital (i.e. computer-readable) format**. ... Digitization is of crucial importance to data processing, storage and transmission, because it "allows information of all kinds in all formats to be carried with the same efficiency and also intermingled".

3. Preprocessing

Image preprocessing are **the steps taken to format images before they are used by model training and inference**. This includes, but is not limited to, resizing, orienting, and color corrections. ... Thus, a transformation that could be an augmentation in some situations may best be a preprocessing step in others.

4. Segmentation

Image segmentation is a branch of digital image processing which focuses on partitioning an image into different parts according to their features and properties. The primary goal of image segmentation is to simplify the image for easier analysis. ... The parts in which you divide the image are called Image Objects.

5. Feature Extraction

Feature extraction refers to the process of transforming raw data into numerical features that can be processed while preserving the information in the original data set. It yields better results than applying machine learning directly to the raw data. ... An example of a simple feature is the mean of a window in a signal.

6. Classification

Classification is a process of categorizing a given set of data into classes, It can be performed on both structured or unstructured data. The process starts with predicting the class of given data points. The classes are often referred to as target, label or categories

7. Interpretation

Interpretation of a machine learning model is the process wherein we try to understand the predictions of a machine learning model

---

2>Explain EM spectrum

When all the electromagnetic radiations are arranged in order of increasing wavelengths, and decreasing frequencies the complete spectrum is called electromagnetic spectrum

Although all electromagnetic waves travel at the speed of light in a vacuum, they do so at a wide range of frequencies, wavelengths, and photon energies.

The various portions has different names based on differences in behaviour in the emission, transmission, and absorption of the corresponding waves and also based on their different practical applications.
There are no precise accepted boundaries between any of these contiguous portions, so the ranges tend to overlap.

The entire electromagnetic spectrum, from the lowest to the highest frequency (longest to shortest wavelength), includes all radio waves (e.g., commercial radio and television, microwaves, radar), infrared radiation, visible light, ultraviolet radiation, X-rays, and gamma rays. Nearly all frequencies and wavelengths of electromagnetic radiation can be used for spectroscopy.

Radio waves in MRI:

Here are some specific Applications:
– Automatic detection of tumors, characterizing their types,
– Measurement of normal/abnormal structures,
– Visualization of anatomy, surgery guidance, therapy planning,
– Exploring relationship between clinical, genomic, and imaging based markers.

¡ Used to produce images of soft tissues, fluid, fat and bone.

¡ Does this by producing a map which depends on the density of hydrogen in the body.
¡ When a person is lying in the magnetic field of the MRI scanner the hydrogen atoms in their body line up.
¡ An radiofrequency field is applied, which causes some of the protons to flip around and the atoms spin together.
¡ This in turn produces a change in magnetic field, which induces a voltage in a coil, and the signal is used to produce an image, dependent upon the tightness of the protons.

3>What is Resolution Compare different Resolution

The term image resolution is applied to digital images, film images, and other types of images and it describes the details that an image holds.
resolution refers to the level of detail to which a ground feature can be described and mapped.
Image resolution can be measured in various ways like spatial, spectral, temporal and radiometric.

Based on these parameters image resolution is categorised into following four types:

• spatial resolution – it refers to variations in the reflectance or emittance determined by the shape, size and texture of the target

• spectral resolution – it refers to the changes in the reflectance or emittance as a function of wavelength

• temporal resolution – it involves seasonal changes in reflectance or emittance

• radiometric resolution – it includes changes in the polarisation of the radiation reflected or emitted by an object.

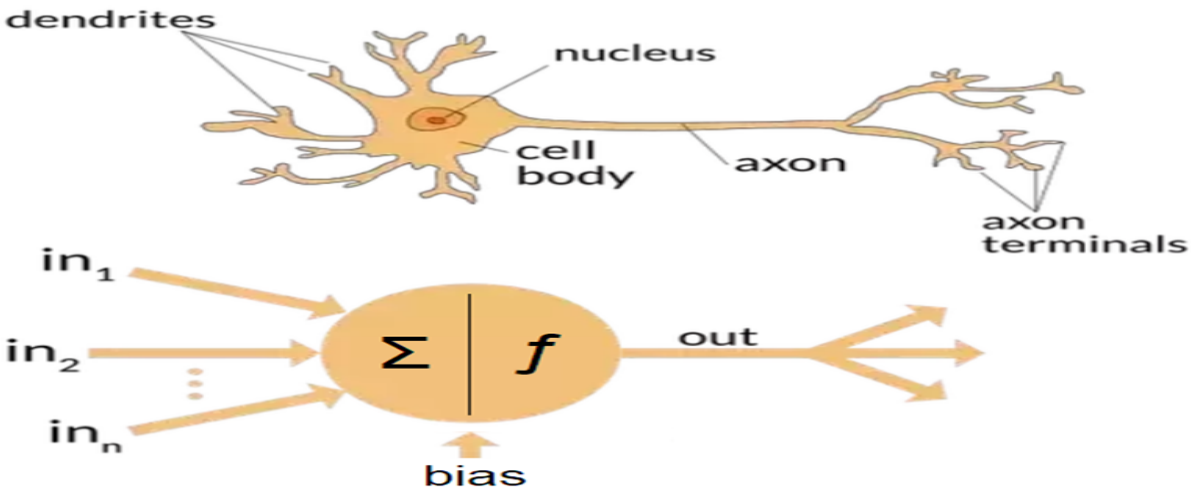
4>Explain DCT, DWT, HOUGH ,transforms

DCT (Discrete cosine transform): The discrete cosine transform (DCT) is used to separate the image in to pixel. DCT is used in signal, image processing especially for lossy compression because it has a strong energy compaction. The lossy image compression ratio of the image was good in number. But the outcome of the image was not good

DWT (Discrete wavelet transforms): Dwt is used to separate the image into a pixel. The DWT represents the two images representing the technique to transform the DWT process. Then the DWT image will move on to the quantization process. That the process is doing again and again to get the best result. Thus the output of the DWT image compression is good. The PSNR value is also good in compression ratio. The quality of the DWT image is also good.

The Hough transform: is a technique which can be used to isolate features of a particular shape within an image. Because it requires that the desired features be specified in some parametric form, the *classical* Hough transform is most commonly used for the detection of regular curves such as lines, circles, ellipses, *etc.* A *generalized* Hough transform can be employed in applications where a simple analytic description of a feature(s) is not possible.

5>Explain Biological Neural Network



Biological Neural Network (BNN) is a structure that consists of Synapse, dendrites, cell body, and axon. In this neural network, the processing is carried out by neurons. Dendrites receive signals from other neurons, Soma sums all the incoming signals and axon transmits the signals to other cells.

Some advantages of BNN :

- The synapses are the input processing element.
- It is able to process highly complex parallel inputs.
- Information is distributed into the network throughout into sub-nodes, even if it gets corrupted it can be retrieved.

Some disadvantages of BNN :

- There is no controlling mechanism.
- Speed of processing is slow being it complex.
- There is no control unit to monitor the information being processed into the network.

6>Compare and contrast AI,DL and ML

1. Artificial Intelligence (AI)

Humans have been obsessed with automation since the beginning of technology adoption. AI enables machines to think without any human intervention. It is a broad area of computer science. AI systems fall into three types: ANI: Artificial Narrow Intelligence, which is goal-oriented and programmed to perform a single task. AGI (Artificial General Intelligence) which allows machines to learn, understand, and act in a way that is indistinguishable from humans in a given situation. ASI (Artificial Super Intelligence) is a hypothetical AI where machines are capable of exhibiting intelligence that surpasses brightest humans.

2. Machine Learning (ML)

ML is a subset of AI that uses statistical learning algorithms to build smart systems. The ML systems can automatically learn and improve without explicitly being programmed. The recommendation systems on music and video streaming services are examples of ML. The machine learning algorithms are classified into three categories: supervised, unsupervised and reinforcement learning.

3. Deep Learning (DL)

This subset of AI is a technique that is inspired by the way a human brain filters information. It is associated with learning from examples. DL systems help a computer model to filter the input data through layers to predict and classify information. Deep Learning processes information in the same manner as the human brain. It is used in technologies such as driver-less cars. DL network architectures are classified into Convolutional Neural Networks, Recurrent Neural Networks, and Recursive Neural Networks.

9>What are Limitations of ANN?

1. Hardware Dependence:
    1. Artificial Neural Networks require processors with parallel processing power, by their structure.
    2. For this reason, the realization of the equipment is dependent.
2. Unexplained functioning of the network:
    1. This is the most important problem of ANN.
    2. When ANN gives a probing solution, it does not give a clue as to why and how.
    3. This reduces trust in the network.
3. Assurance of proper network structure:
    1. There is no specific rule for determining the structure of artificial neural networks.
    2. The appropriate network structure is achieved through experience and trial and error.
4. The difficulty of showing the problem to the network:
    1. ANNs can work with numerical information.
    2. Problems have to be translated into numerical values before being introduced to ANN.
    3. The display mechanism to be determined will directly influence the performance of the network.
5. The duration of the network is unknown:

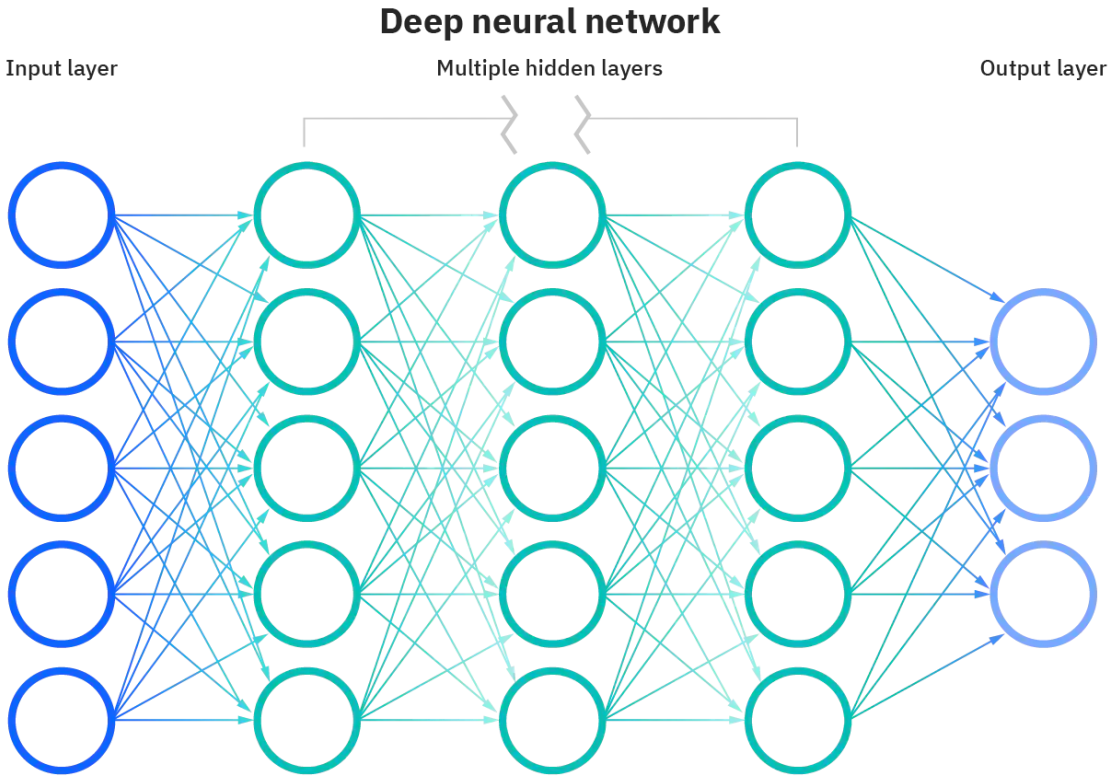
1. The network is reduced to a certain value of the error on the sample means that the training has been completed.
2. The value does not give us optimum results.

10>Emphasize on Deep Learning.

Ans) Deep Learning, is a type of Machine Learning, inspired by the structure of a human brain. Deep learning algorithms attempt to draw similar conclusions as humans would by continually analyzing data with a given logical structure. To achieve this, deep learning uses a multi-layered structure of algorithms called neural networks.

The individual layers of neural networks can also be thought of as a sort of filter that works from gross to subtle, increasing the likelihood of detecting and outputting a correct result.

The human brain works similarly. Whenever we receive new information, the brain tries to compare it with known objects. The same concept is also used by deep neural networks.



11>Compare image enhancement and image restoration.

Sr. No.	Image Enhancement	Image Restoration
1.	As the name suggests, in Image Enhancement, the original image is processed so that the resultant image is more suitable than the original for specific applications.	The aim of image restoration is to bring the image towards what it would have been if it had been recorded without degradation.
2.	Image enhancement makes a picture look better, without regard to how it really truly should look.	Image restoration tries to fix the image to get back to the real, true image.
3.	Image enhancement means improving the image to show some hidden details.	Image restoration means improving the image to match the original image.
4.	Image enhancement is a purely subjective processing technique.	Image restoration is an objective process.
5.	Image enhancement is a cosmetic procedure i.e. it does not add any extra information to the original image. It merely improves the subjective quality of the images by work in with the existing data.	Restoration tries to reconstruct by using a priori knowledge of the degradation phenomena. Restoration hence deals with getting an optimal estimate of the desired result

12>Explain sampling and quantization.

In Digital Image Processing, signals captured from the physical world need to be translated into digital form by "Digitization" Process. In order to become suitable for digital processing, an image function f(x,y) must be digitized both spatially and in amplitude.

Digitization process involves two main processes called

1. Sampling: Digitizing the co-ordinate value is called sampling.
2. Quantization: Digitizing the amplitude value is called quantization

Sampling:
Since an analogue image is continuous not just in its co-ordinates (x axis), but also in its amplitude (y axis), so the part that deals with the digitizing of coordinates is known as sampling. In digitizing sampling is done on independent variables. In the case of equation y = sin(x), it is done on the x variable.

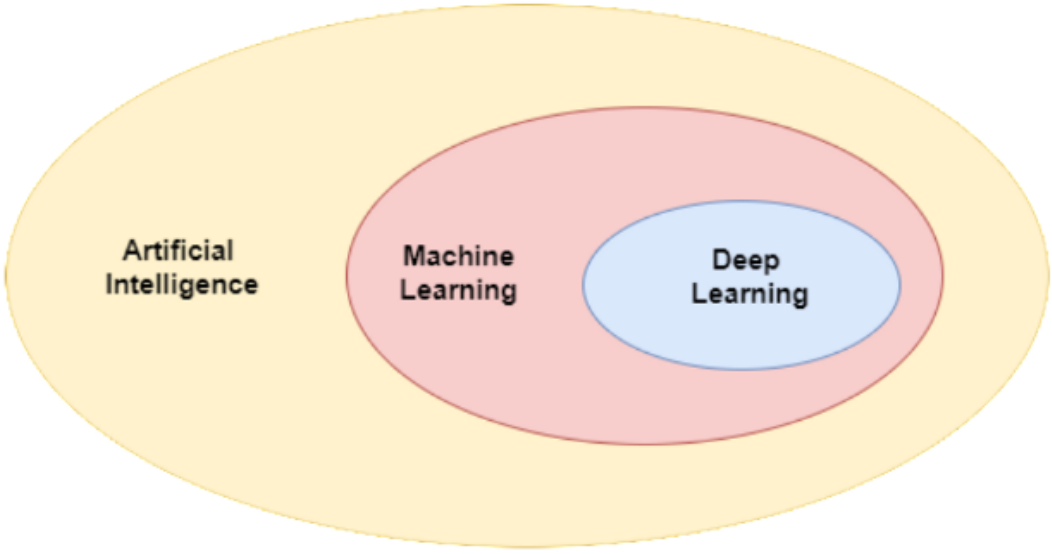
When looking at this image, we can see there are some random variations in the signal caused by noise. In sampling we reduce this noise by taking samples. It is obvious that the more samples we take, the quality of the image would be better, the noise would be more removed and the same happens vice versa. However, if you take sampling on the x axis, the signal is not converted to digital format, unless you take sampling of the y-axis too which is known as quantization.

Quantization
It is opposite to sampling because it is done on “y axis” while sampling is done on “x axis”. Quantization is a process of transforming a real valued sampled image to one taking only a finite number of distinct values. Under quantization process the amplitude values of the image are digitized. In simple words, when you are quantizing an image, you are actually dividing a signal into quanta(partitions).

UNIT 2

2,3,4,5> AI vs ML vs DL & Supervised vs Unsupervised vs Reenforced learning And CNN RNN Architecture & CNN vs RNN

Artificial Intelligence, Machine Learning, Deep Learning, Data Science are popular terms in this era. And knowing what it is and the difference between them is more crucial than ever. Although these terms might be closely related there are differences between them see the image below to visualize it.



What is Artificial Intelligence?

Humans have long been obsessed with creating AI ever since the question, “Can machines think?”, was posed by Alan Turing in 1950. AI enables the machine to think, that is without any human intervention the machine will be able to take its own decision. It is a broad area of computer science that makes machines seem like they have human intelligence. So it’s not only programming a computer to drive a car by obeying traffic signals but it’s when that program also learns to exhibit the signs of human-like road rage.

Types of Artificial Intelligence System

AI systems are classified by their ability to imitate human behaviors, the hardware they use to do so, their applications in the real world and the theory of mind. Using these features for comparison, all systems of artificial intelligence actual and hypothetical fall into one of three types:

ANI: Artificial Narrow Intelligence

Artificial Narrow intelligence is also known as weak AI and it is the only type of AI that exists in our world today. Narrow AI is goal oriented and is programmed to perform a single task and is very intelligent in completing the specific task that it is programmed to do. Some examples of ANI are Siri, Auto pilot in an airplane, chat bots, self driving cars etc.

Narrow AI systems are not conscious, sentient or driven by emotions as humans are, they use information from a specific dataset and do not perform any task that is outside of the single task that they are designed to perform.

AGI: Artificial General Intelligence

Artificial General Intelligence also referred to as strong AI is a concept in which machines exhibit human intelligence. In this the machines have the ability to learn, understand and act in a way that is indistinguishable from a human in a given situation. The General AI does not currently exist but has been used in many sci-fi hollywood movies in which the humans interact with machines that are conscious, driven by emotions and self-aware.

Using strong AI we can have the ability to build machines that can think, strategize and perform multiple tasks under uncertain conditions. They can integrate their prior knowledge in decision making to come up with innovative, creative and unconventional solutions.

ASI: Artificial Super Intelligence

I am sure you remember Arnold Schwarzenegger's "The Terminator" where a machines cognizance superseded human intelligence in all aspects. Artificial Super Intelligence is a hypothetical AI where machines will be capable of exhibiting intelligence that surpasses that of the brightest humans. In this type of AI, apart from having multifaceted intelligence of human beings machines will have greater problem solving and decision making capabilities that will be far superior than human beings. It is the type of AI that will have great impact on humanity and may lead to extinction of the human race from the planet.

What is Machine Learning?

Machine Learning is a subset of Artificial Intelligence that uses statistical learning algorithms to build systems that have the ability to automatically learn and improve from experiences without being explicitly programmed.

Most of us use machine learning in our day to day life when we use services like recommendation systems on Netflix, Youtube, Spotify; search engines like

google and yahoo; voice assistants like google home and amazon alexa. In Machine Learning we train the algorithm by providing it with a lot of data and allowing it to learn more about the processed information.

ML algorithms can be broadly classified into three categories Supervised, Unsupervised and Reinforcement learning.

Supervised Learning

In supervised learning we have input variables (x) and an output variable (Y) and we use an algorithm to learn the mapping from input to output. In other words, a supervised learning algorithm takes a known set of input dataset and its known responses to the data (output) to learn the regression/classification model. A learning algorithm then trains a model to generate a prediction for the response to new data or the test datasets.

Unsupervised Learning

Unsupervised Learning is used when we do not have labelled data. Its main focus is to learn more about the data by inferring patterns in the dataset without reference to the known outputs. It is called unsupervised because the algorithms are left on their own to group the unsorted information by finding similarities, differences and patterns in the data. Unsupervised learning is mostly performed as a part of exploratory data analysis. It is most commonly used to find clusters of data and for dimensionality reduction.

Reinforcement Learning

In simple terms, reinforcement learning can be explained as learning by continuously interacting with the environment. It is a type of machine learning algorithm in which an agent learns from an interactive environment in a trial and error way by continuously using feedback from its previous actions and experiences. The reinforcement learning uses rewards and punishments, the agents receive rewards for performing correct actions and penalties for doing it incorrectly.

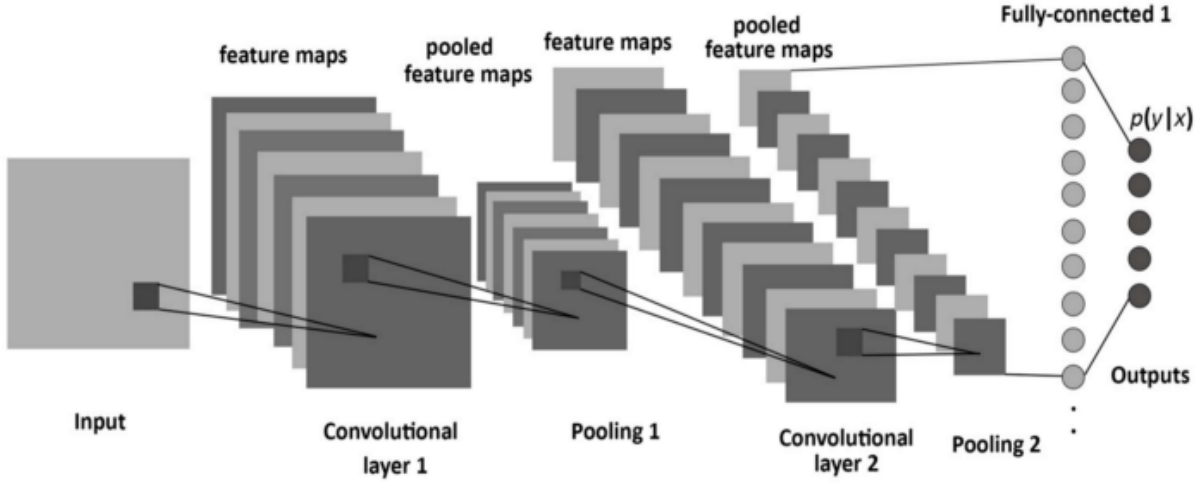
What is Deep Learning?

Deep learning is a machine learning technique that is inspired by the way a human brain filters information, it is basically learning from examples. It helps a computer model to filter the input data through layers to predict and classify information. Since deep learning processes information in a similar manner as a human brain does, it is mostly used in applications that people generally do. It

is the key technology behind driver-less cars, that enables them to recognize a stop sign and to distinguish between a pedestrian and lamp post. Most of the deep learning methods use neural network architectures, so they are often referred to as deep neural networks.

Deep Learning is basically mimicking the human brain, it can also be defined as a multi neural network architecture containing a large number of parameters and layers. The three fundamental network architectures are as listed below:

Convolutional Neural Networks



Convolutional Neural Network is basically an artificial neural network that is most widely used in the field of Computer Vision for analyzing and classifying images. It is a deep learning algorithm that takes the input image and assigns weights/biases to various aspects or objects in the image, so that it can differentiate one from the other. The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers, and normalization layers. The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex.

CNN Architecture (Unit 2)

An input layer, an output layer, and multiple hidden layers make up convolutional networks. The neurons in the layers of a convolutional network are arranged in three dimensions, unlike those in a standard neural network (width, height, and depth dimensions). This enables the CNN to convert a three-dimensional input volume into an output volume. Convolution, pooling, normalizing, and fully connected layers

make up the hidden layers. Multiple convolutional layers are used in CNNs to filter input volumes to higher levels of abstraction.

Recurrent Neural Networks

Recurrent Neural Networks is a type of neural network architecture that is used in sequence prediction problems and is heavily used in the field of Natural Language Processing. RNNs are called *recurrent* because they perform the same task for every element of a sequence, with the output being dependent
on the previous computations. Another way to think about RNNs is that they have a "memory" which captures information about what has been calculated so far.

Convolutional Neural Networks	Recurrent Neural Networks
In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most commonly applied to analyzing visual imagery.	A recurrent neural network (RNN) is a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence.
It is suitable for spatial data like images.	RNN is used for temporal data, also called sequential data.
CNN is a type of feed-forward artificial neural network with variations of multilayer perceptron's designed to use minimal amounts of preprocessing.	RNN, unlike feed-forward neural networks- can use their internal memory to process arbitrary sequences of inputs.
CNN is considered to be more powerful than RNN.	RNN includes less feature compatibility when compared to CNN.
This CNN takes inputs of fixed sizes and generates fixed size outputs.	RNN can handle arbitrary input/output lengths.
CNN's are ideal for images and video processing.	RNNs are ideal for text and speech analysis.
Applications include Image Recognition, Image Classification, Medical Image Analysis, Face Detection and Computer Vision.	Applications include Text Translation, Natural Language Processing, Language Translation, Sentiment Analysis and Speech Analysis.

ANN

Features of ANN:

(i) Artificial neural networks are extremely powerful computational devices

(ii) ANNs are modeled on the basis of current brain theories, in which information is represented by weights.

(iii) ANNs have massive parallelism which makes them very efficient.

(iv) They can learn and generalize from training data so there is no need for enormous feats of programming.

(v)Storage is fault tolerant i.e. some portions of the neural net can be removed and there will be only a small degradation in the quality of stored data.

(vi) They are particularly fault tolerant which is equivalent to the "graceful degradation" found in biological systems.

(vi) Data are naturally stored in the form of associative memory which contrasts with conventional memory, in which data are recalled by specifying address of that data.

(viii) They are very noise tolerant, so they can cope with situations where normal symbolic systems would have difficulty.

(ix) In practice, they can do anything a symbolic/ logic system can do and more.

(x) Neural networks can extrapolate and intrapolate from their stored information. The neural networks can also be trained. Special training teaches the net to look for significant features or relationships of data.

Uses of ANN:

(i) Image processing and character recognition.

(ii) Forcasting.

Advantages

- The main advantage of ANN is parallel processing. This makes it more useful in linear programs.
- Due to their parallel processing structure, any failure in one neural element will not affect the rest of the process.

- Neural networks can be applied to any application and they can solve any complex problem.
- By implementing appropriate learning algorithms, an ANN can be made to learn without reprogramming.

Disadvantages

- All the parallel processing requires a huge amount of processing power and time.
- There is a requirement for a “training” period before real-world implementation.

Activation Functions

Name	Plot	Equation	Derivative
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic (a.k.a Soft step)		$f(x) = \frac{1}{1+e^{-x}}$	$f'(x) = f(x)(1-f(x))$
TanH		$f(x) = \tanh(x) = \frac{2}{1+e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
ArcTan		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2+1}$
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parameteric Rectified Linear Unit (PReLU) [2]		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Exponential Linear Unit (ELU) [3]		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1+e^x)$	$f'(x) = \frac{1}{1+e^{-x}}$

Rectified Linear Unit (ReLU) Function

One of the most popular AFs in DL models, the rectified linear unit (ReLU) function, is a fast-learning AF that promises to deliver state-of-the-art performance with stellar results. Compared to other AFs like the sigmoid and tanh functions, the ReLU function offers much better performance and generalization in deep learning. The function is a nearly linear function that retains the properties of linear models, which makes them easy to optimize with gradient-descent methods.

The ReLU function performs a threshold operation on each input element where all values less than zero are set to zero. Thus, the ReLU is represented as:

$$f(x) = max(0, x) = \begin{cases} x_i, & if\ x_i \geq 0 \\ 0, & if\ x_i < 0 \end{cases} \quad -(1.14)$$

7>Compare forward propagation and back propagation in ANN.

Forward Propagation: This is the process by means of which a neural network takes input data and keeps on producing another value, which is fed into the subsequent layer of neural network. This implies that the network is 'Connected'. Finally, the output is put into a loss function and used to find the loss. Here, ends your forward propagation.

Backward Propagation: This is the process wherein, gradients (derivative of loss computed w.r.t. weights/parameters of neural network) are computed. First, the gradient of parameters of last layer is computed. Now, this computed gradient is used to compute the gradients of penultimate layer and so on. Basically in back propagation, information flows from right to left (output layer to input layer) as against left to right (input to output layer) in forward propagation.

When gradients of a particular layer are found, its weights are also simultaneously updated.

UNIT 3 & UNIT 4

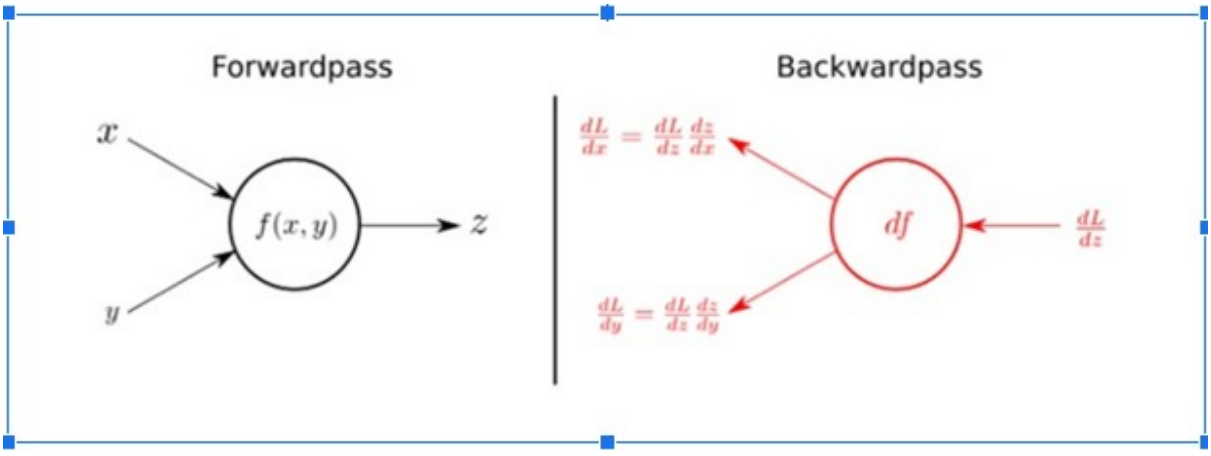
1. Explain forward and backward propagation in Neural Network? Consider the Radiological CT/MRI diseases classification.

Forward Propagation - As the name suggests, the input data is fed in the forward direction through the network.
At each neuron in a hidden or output layer, the processing happens in two steps:

1. **Pre-activation**: it is a weighted sum of inputs i.e., the linear transformation of weights w.r.t to inputs available. Based on this aggregated sum and activation function the neuron makes a decision whether to pass this information further or not.
2. **Activation**: the calculated weighted sum of inputs is passed to the activation function. An activation function is a mathematical function which adds non-linearity to the network. There are four commonly used and popular activation functions — sigmoid, hyperbolic tangent(tanh), ReLU and Softmax.

Backward Propagation

Back-propagation is the essence of neural net training. It is the practice of fine-tuning the weights of a neural net based on the error rate (i.e., loss) obtained in the previous epoch (i.e., iteration). Proper tuning of the weights ensures lower error rates, making the model reliable by increasing its generalization.



2. Explain Limitations of Neural Network that made to evolve Deep Neural Networks? Draw and emphasis on Deep Neural Network Architecture, explain each block? Explain with example, challenges in Deep Neural Network optimization?

3.Consider at least 8 activation functions and explain advantages and disadvantages of each.

Function	**Advantages**	**Disadvantages**
Sigmoid	Gives you a smooth gradient while converging. Gives a clear prediction with 1 & 0.	Prone to Vanishing Gradient problem. Not a zero-centric function (Always gives a positive value).
Tanh(Hyperbolic Tangent Movement):-	It is Normally used as the input of a binary probabilistic function. Zero-centric function Gives you a smooth gradient while converging..	Prone to Vanishing Gradient function. Computationally function(exponential in nature).
RELU:- (Rectified Linear Unit)	Can deal with Vanishing Gradient problem. Computationally inexpensive	Not a zero-centric function. Gives zero value, as

	in nature).	inactive in the negative axis.
SWISH	Can deal with Vanishing Gradient problem. Also known as self gated activation function is one of the kinds that is being inspired by the use of the Sigmoid function inside an LSTM(Long Short based network. The output is a workaround and Sigmoid function which helps in normalising	Computationally expensive function (as of Sigmoid).
SoftPlus	Convergence of gradient is RELU function. It can handle the Vanishing Gradient problem.	Computationally expensive than RELU(as exponential in nature).

Explain in General Image Classification using CNN? What is a pre-trained CNN model?

1. **Choose a Dataset**

Choose a dataset of your interest or you can also create your own image dataset for solving your own image classification problem

2. **Prepare Dataset for Training.**

Preparing our dataset for training will involve assigning paths and creating categories (labels), resizing our images.

3. **Create Training Data**

Training is an array that will contain image pixel values and the index at which the image in the CATEGORIES list.

4. **Shuffle the Dataset**

Shuffle ur dataset in Relevant Manner

5. **Assigning Labels and Features**

This shape of both the lists will be used in Classification using the NEURAL NETWORKS.

6. Normalising X and converting labels to categorical data
7. Split X and Y for use in CNN
8. Define, compile and train the CNN Model
9. pre-trained CNN model

A pre-trained model is a model created and trained by someone else to solve a problem that is similar to ours. In practice, someone is almost always a tech giant or a group of star researchers. They usually choose a very large dataset as their base datasets such as ImageNet or the Wikipedia Corpus.

➔ With relevant details explain each stage of CNN?

- Step 1: Convolution Operation

The term convolution refers to **the mathematical combination of two functions to produce a third function**. It merges two sets of information. In the case of a CNN, the convolution is performed on the input data with the use of a filter or kernel (these terms are used interchangeably) to then produce a feature map

Step 1(b): ReLU Layer

The ReLU layer applies the **function f(x) = max(0, x)** to all of the values in the input volume. In basic terms, this layer just changes all the negative activations to 0. This layer increases the nonlinear properties of the model and the overall network without affecting the receptive fields of the conv layer.

Step 2: Pooling

Pooling layers are **used to reduce the dimensions of the feature maps**. Thus, it

reduces the number of parameters to learn and the amount of computation performed in the network. The pooling layer summarises the features present in a region of the feature map generated by a convolution layer.

Step 3: Flattening

- Flattening is **converting the data into a 1-dimensional array** for inputting it to the next layer. We flatten the output of the convolutional layers to create a single long feature vector. And it is connected to the final classification model, which is called a fully-connected layer.
- Step 4: Full Connection

Fully Connected Layer is simply, **feed forward neural networks**. Fully Connected Layers form the last few layers in the network. The input to the fully connected layer is the output from the final Pooling or Convolutional Layer, which is flattened and then fed into the fully connected layer

---

➔ Explain Binary and Multi Labelled classification using CNN model? Explain any simple CNN pretrained model with examples. (ResNet-34 and ResNet-50)

Binary classification is used in the machine learning domain commonly. It is the simplest way to classify the input into one of the two possible categories. For example, give the attributes of apple-like Color, weight, etc. that classify the fruit as either a green apple or red apple.

Convolution Neural Network(CNN). It is a class of Neural network that has proven very effective in areas of image recognition, processing, and classification.

CNN model requires training data for training weights and validation for checking its performance. Each input images passes through a series of convolution layers with filters(Kernels) :

1. Convolution layers
2. Pooling layers
3. Fully-connected layer (FC) applying the sigmoid function.

the sigmoid function is used to classify an object with a probabilistic value which turns out as 0 or 1 for binary classification.

the Convolution + MaxPooling layers act as feature extractors from the input image while a fully connected layer acts as a classifier. Oreceiving an image as input, the network will assign assigns the highest probability for it and predict which class the input image belongs to.

In multi-label classification, one data sample can belong to multiple classes(labels).

- Predicting movie genre from a movie poster is an example of multi-label classification, where a movie can have multiple genres.

- We use the sigmoid activation function on the final layer. Sigmoid converts each score of the final node between 0 to 1 independent of what the other scores are.

- If the score for some class is more than 0.5, the data is classified into that class. And there could be multiple classes having a score of more than 0.5 independently. Thus the data could be classified into multiple classes.

➔ What is Image Segmentation? With Detailed diagram explain U-Net Segmentation? Emphasis on its limitation.

Image segmentation is the process of partitioning a digital image into multiple **image segments**, also known as **image regions** or **image objects** (sets of pixels).

U-Net is an architecture for semantic segmentation. It consists of a contracting path and an expansive path. The contracting path follows the typical architecture of a convolutional network. It consists of the repeated application of two 3x3 convolutions (unpadded convolutions), each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for downsampling. At each downsampling step we double the number of feature channels. Every step in the expansive path consists of an upsampling of the feature map followed by a 2x2 convolution (“up-convolution”) that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU.

The architecture consist of two paths -

1. First, The encoder path — The left hand side of the ‘U’ Shaped architecture is called encoder/contraction path. This path consist of stack of convolutional layer and max pooling layer. This is used to capture the context of the image.
2. Second, The decoder path — The right hand side of the ‘U’ shaped architecture is called decoder/expansion path. The path consist of transposed convolutional layers. This is used to expand the enable precise localization.

Limitations

(1) their optimal depth is unknown, requiring extensive architecture search or inefficient ensemble of models of varying depths.

(2) their skip connections impose an unnecessarily restrictive fusion scheme, forcing aggregation only at the same-scale feature maps of the encoder and decoder sub-networks

---

➔ **Which all paraments are considered for performance analysis of CNN classification and U- Net Segmentation? (Evaluation Metrics)**

➔ What is feature extraction, Explain How CNN automatically extract features of input image and classify Image?

Feature extraction is the **process of defining a set of features, or image characteristics**, which will most efficiently or meaningfully represent the information that is important for analysis and classification.

CNN is a neural network that extracts input image features and another neural network classifies the image features.

The input image is used by the feature extraction network. The extracted feature signals are utilized by the neural network for classification.

The neural network classification then works on the basis of the image features and produces the output.

The neural network for feature extraction includes convolution layer piles and sets of pooling layers.

As its name implies, the convolution layer transforms the image using the process of the convolution.

It can be described as a series of digital filters. The layer of pooling transforms the neighboring pixels into a single pixel.

The pooling layer then decreases the image dimension.

➔ What makes CNN to be powerful model in Image Recognition and classification applications?