Representations, Description and Recognition

Image Processing | CSIT 5^{th} Sem

Yuba Raj Devkota | NCCS

Unit 7	**Representations, Description and Recognition**	**Teaching Hours (5)**
Representation and Descriptions	Introduction to some descriptors: Chain codes, Signatures, Shape Numbers, Fourier Descriptors	2 hrs.
Recognition	Patterns and pattern classes, Decision-Theoretic Methods, Introduction to Neural Networks and Neural Network based Image Recognition	2 hrs.
Pattern Recognition	Overview of Pattern Recognition with block diagram	1 hr

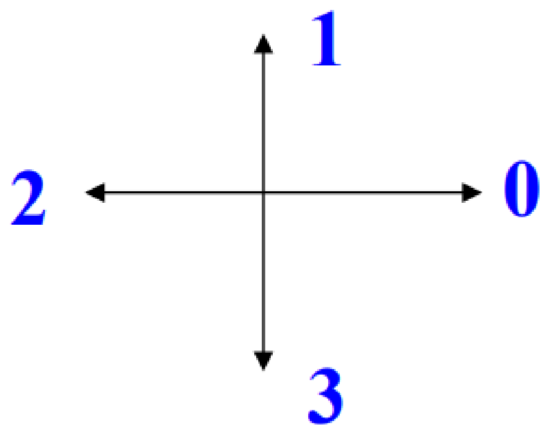
Image Representation & Description

- After an image is segmented into regions, the regions are represented and described in a form suitable for computer processing (descriptors).
- Representing a region:
  1. In terms of its external characteristics (boundary)
  2. In term of its internal characteristics

  Exp: A region might be represented by the length of its boundary.
  - External representations are used when the focus is on shape of the region.
  - Internal representations are used when the focus is on color and texture.
  - Representations should be insensitive to rotation and translation.

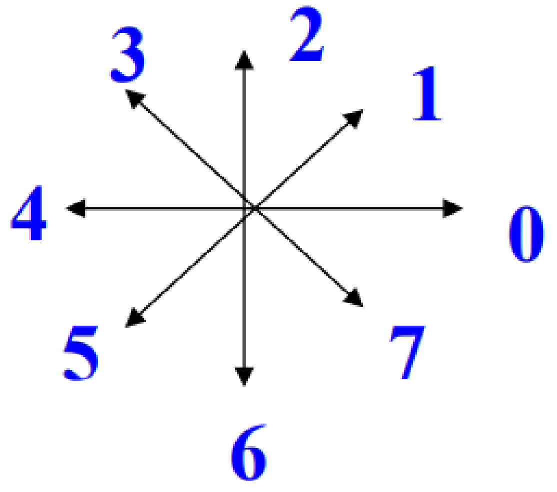
Chain Code

- Chain Code: Used to represent a boundary by a connected sequence of straight line segments.
  - 4 or 8 connectivity is used
  - The direction of each segment is coded by a numbering scheme.



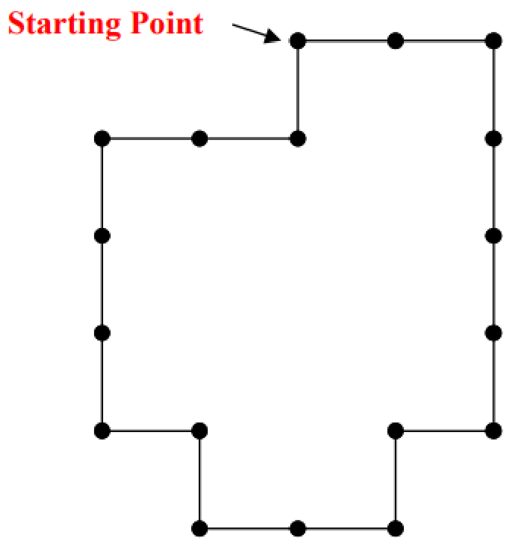
Method:

- Follow the boundary in a specific (clockwise) direction.
- Assign a direction to the segment connecting every pair of pixels.



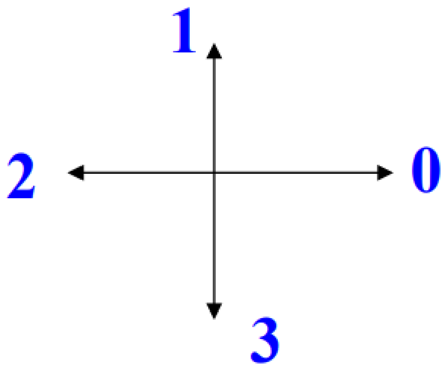
Exp: 00333323221211 1001

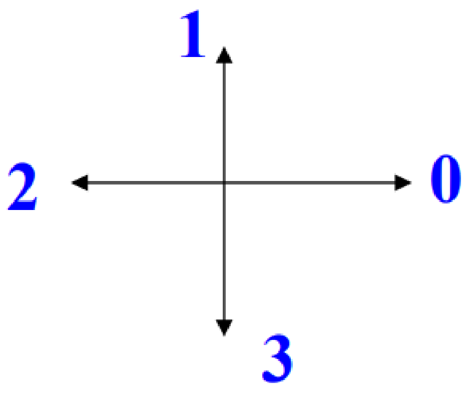
- Problems:
  - The chain code depends on the starting point.
  - It changes with rotation.



Solutions:

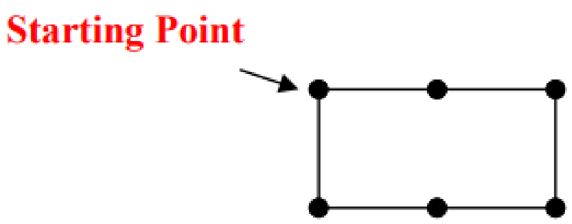
- Treat the chain code as a circular sequence of numbers. Circulate until the number is of minimum magnitude.
- Use the difference of chain code instead of the code itself: count counterclockwise the number of directions that separates two adjacent elements (First difference)



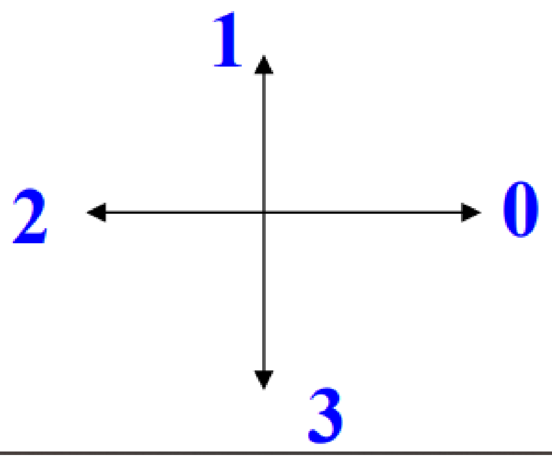


- Exp: 10103322
- First difference code: 33133030

Shape Number

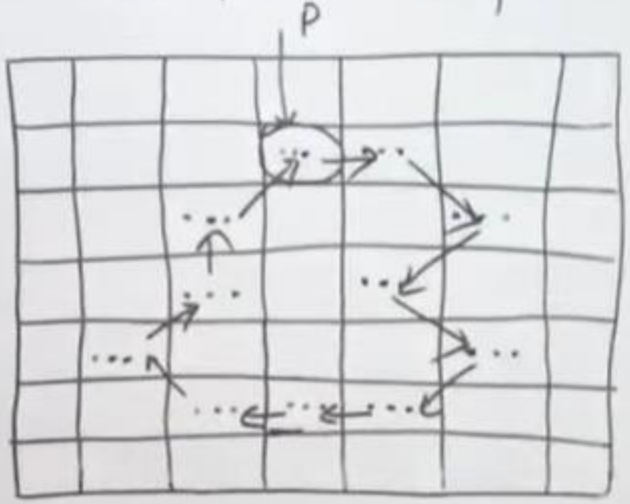


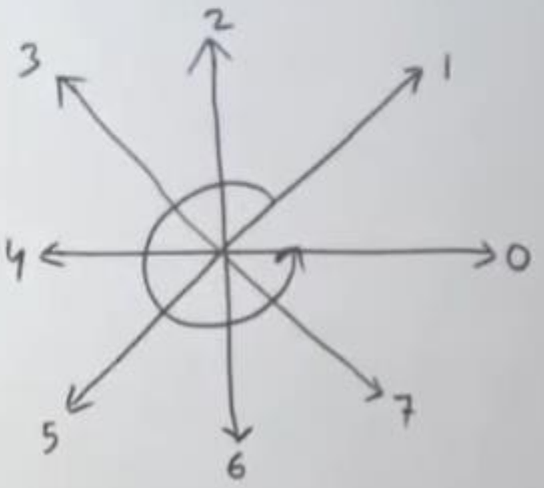
- Shape number: first difference of smallest magnitude (in the chain code)
  - Exp: chain code 003221
  - First difference code: 303303
  - Shape number: 033033



AKTU 2012-13

Q. Given an image, write down the 8-chain code and find shape number of it.





8-chain Co

1) Chain Code - 0 7 5 7 5 4 4 3 1 2 1

First-difference - 7 7 6 2 6 7 0 7 6 1 7

2) Shape no. - 0 7 6 1 7 7 7 6 2 6 7

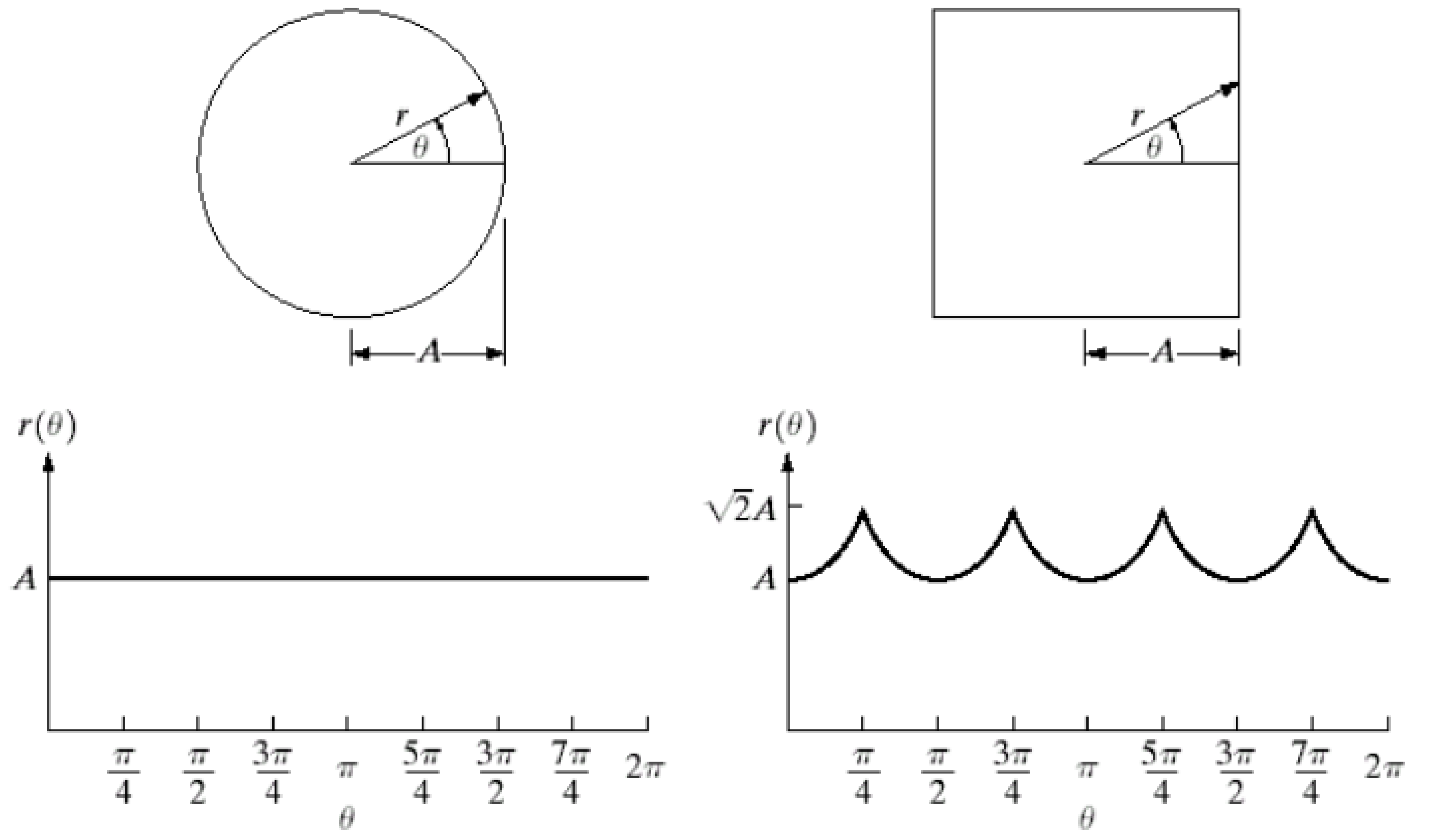
Signature

- Signature: a 1-D functional representation of a boundary
- Different ways of generating signature
- Plot distance form centroid to boundary as a function of angle

- Invariant to translation, depend on rotation and scaling
- To make it invariant to rotation we should select the same starting point regardless of the orientation
  - Select starting point farthest from centroid (if unique)
- To make it invariant to scaling we can normalize to a particular range
- Other signatures: traverse the boundary, at each point plot the angle between a line tangent to the boundary and a reference line
- Slope-density-function: histogram of tangent-angle values
- Straight segments will form the peaks of histogram

a b

FIGURE 11.5 Distance-versus-angle signatures. In (a) $r(\theta)$ is constant. In (b), the signature consists of repetitions of the pattern $r(\theta) = A\sec\theta$ for $0 \le \theta \le \pi/4$ and $r(\theta) = A\csc\theta$ for $\pi/4 < \theta \le \pi/2$.



Fourier Descriptor

N point boundary

$(x_0, y_0), (x_1, y_1), \ldots. (x_{N-1}, y_{N-1})$

$s(k) = x_k + jy_k$

N point DFT of s(k):

$$a(u) = \frac{1}{N}\sum_{k=0}^{N-1} s(k)\exp(-j2\pi uk/N)$$

$a(u)$ are called Fourier descriptors.

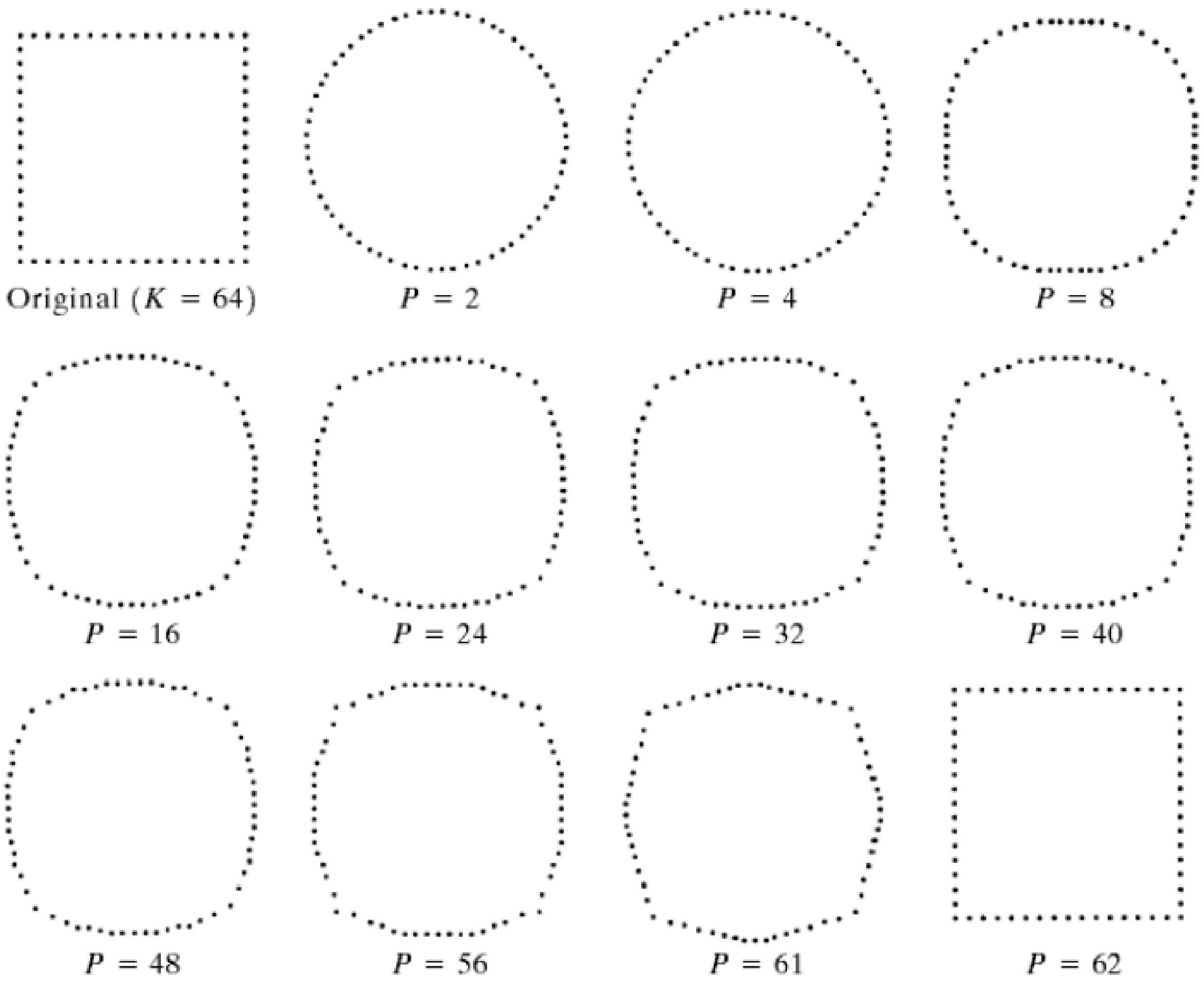
If P of the Fourier descrptors are used

$$\hat{s}(k) = \sum_{u=0}^{P-1} a(u)\exp(j2\pi uk/N)$$

$P < N \Rightarrow$ High frequency details of the boundary (e.g., corners) are removed.

- Fourier descriptors are not directly insensitive to translation, rotation and scaling.
- Magnitude of the Fourier descriptors is insensitive to rotation.

Fourier Descriptor



Pattern Recognition in Image Processing

- **Pattern** is everything around in this digital world. A pattern can either be seen physically or it can be observed mathematically by applying algorithms.
- **Example:** The colors on the clothes, speech pattern, etc. In computer science, a pattern is represented using vector feature values.

What is Pattern Recognition?

- **Pattern recognition** is the process of recognizing patterns by using a machine learning algorithm. Pattern recognition can be defined as the classification of data based on knowledge already gained or on statistical information extracted from patterns and/or their representation. One of the important aspects of pattern recognition is its application potential.
- **Examples:** Speech recognition, speaker identification, multimedia document recognition (MDR), automatic medical diagnosis.

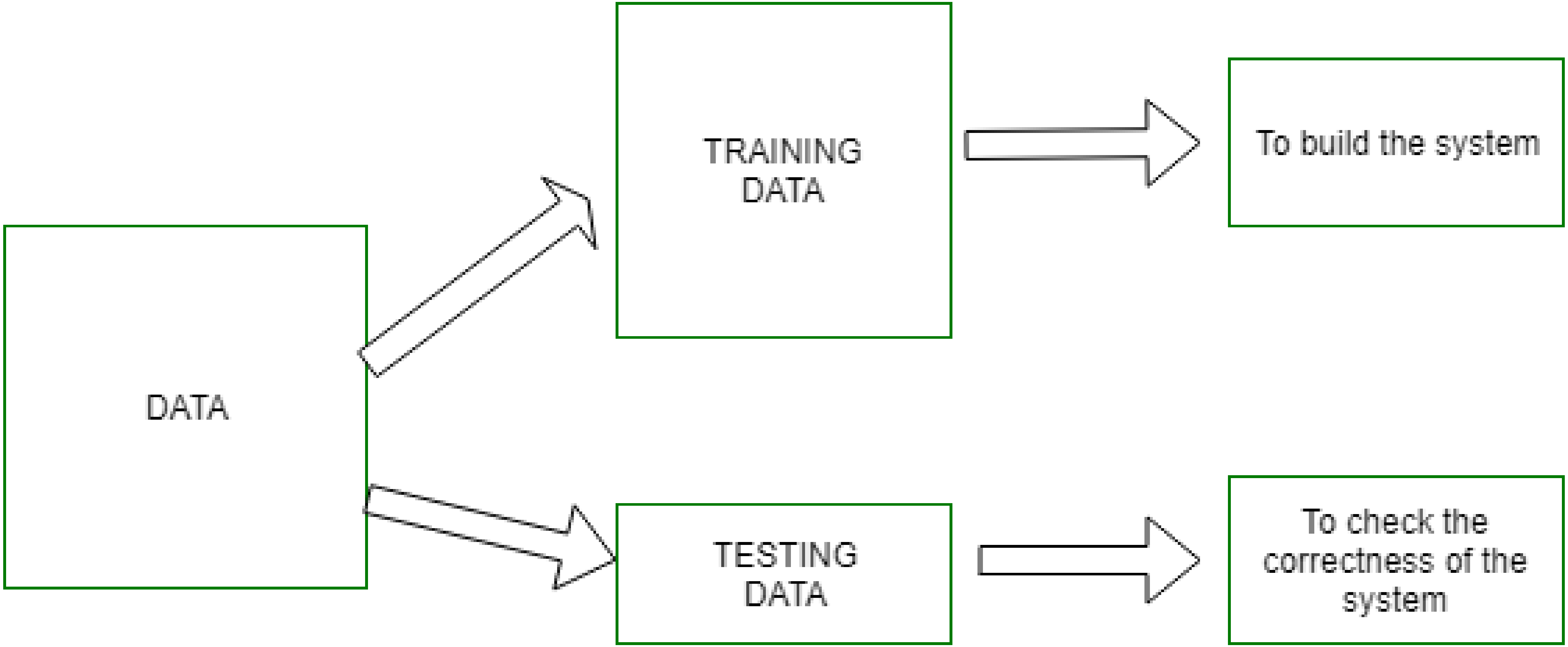
- In a typical pattern recognition application, the raw data is processed and converted into a form that is amenable for a machine to use. Pattern recognition involves the classification and cluster of patterns.
    - In classification, an appropriate class label is assigned to a pattern based on an abstraction that is generated using a set of training patterns or domain knowledge. Classification is used in supervised learning.
    - Clustering generated a partition of the data which helps decision making, the specific decision-making activity of interest to us. Clustering is used in unsupervised learning.
- **Features** may be represented as continuous, discrete, or discrete binary variables. A feature is a function of one or more measurements, computed so that it quantifies some significant characteristics of the object.
- **Example:** consider our face then eyes, ears, nose, etc are features of the face.

- A set of features that are taken together, forms the **features vector**.
- **Example:** In the above example of a face, if all the features (eyes, ears, nose, etc) are taken together then the sequence is a feature vector([eyes, ears, nose]). The feature vector is the sequence of a feature represented as a d-dimensional column vector. In the case of speech, MFCC (Mel-frequency Cepstral Coefficient) is the spectral feature of the speech. The sequence of the first 13 features forms a feature vector.

- **Pattern recognition possesses the following features:**
  - Pattern recognition system should recognize familiar patterns quickly and accurate
  - Recognize and classify unfamiliar objects
  - Accurately recognize shapes and objects from different angles
  - Identify patterns and objects even when partly hidden
  - Recognize patterns quickly with ease, and with automaticity.

Training and Learning in Pattern Recognition

- **Learning** is a phenomenon through which a system gets trained and becomes adaptable to give results in an accurate manner. Learning is the most important phase as to how well the system performs on the data provided to the system depends on which algorithms are used on the data. The entire dataset is divided into two categories, one which is used in training the model i.e. Training set, and the other that is used in testing the model after training, i.e. Testing set.
- **Training set:**
  The training set is used to build a model. It consists of the set of images that are used to train the system. Training rules and algorithms are used to give relevant information on how to associate input data with output decisions. The system is trained by applying these algorithms to the dataset, all the relevant information is extracted from the data, and results are obtained. Generally, 80% of the data of the dataset is taken for training data.
- **Testing set:**
  Testing data is used to test the system. It is the set of data that is used to verify whether the system is producing the correct output after being trained or not. Generally, 20% of the data of the dataset is used for testing. Testing data is used to measure the accuracy of the system. For example, a system that identifies which category a particular flower belongs to is able to identify seven categories of flowers correctly out of ten and the rest of others wrong, then the accuracy is 70 %


DATA
TRAINING DATA
TESTING DATA
To build the system
To check the correctness of the system

- **Real-time Examples and Explanations:**
  A pattern is a physical object or an abstract notion. While talking about the classes of animals, a description of an animal would be a pattern. While talking about various types of balls, then a description of a ball is a pattern. In the case balls considered as pattern, the classes could be football, cricket ball, table tennis ball, etc. Given a new pattern, the class of the pattern is to be determined. The choice of attributes and representation of patterns is a very important step in pattern classification. A good representation is one that makes use of discriminating attributes and also reduces the computational burden in pattern classification.
- An obvious representation of a pattern will be a **vector**. Each element of the vector can represent one attribute of the pattern. The first element of the vector will contain the value of the first attribute for the pattern being considered.

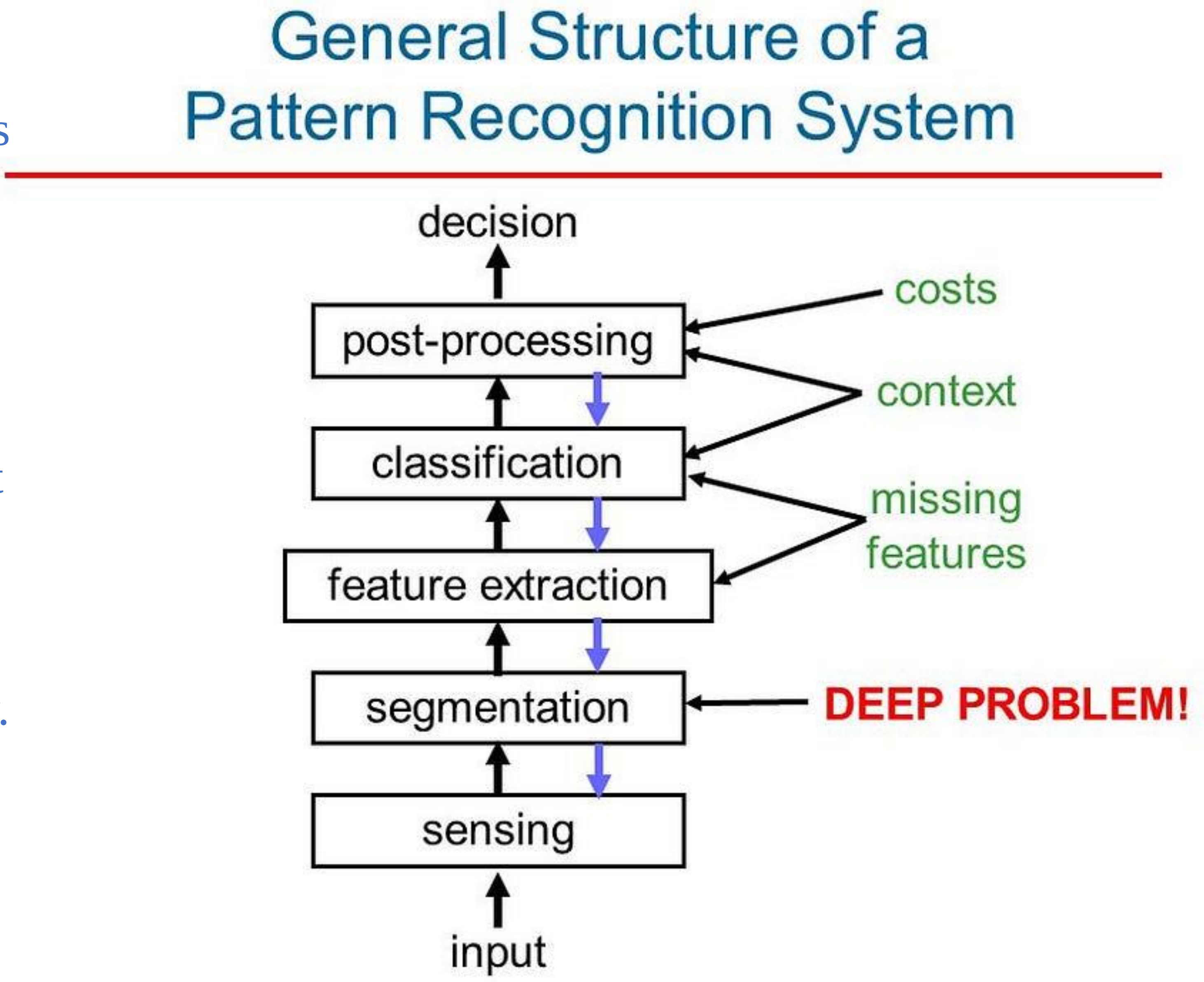
- **Example:** While representing spherical objects, (25, 1) may be represented as a spherical object with 25 units of weight and 1 unit diameter. The class label can form a part of the vector. If spherical objects belong to class 1, the vector would be (25, 1, 1), where the first element represents the weight of the object, the second element, the diameter of the object and the third element represents the class of the object.

- **Advantages:**
- Pattern recognition solves classification problems
- Pattern recognition solves the problem of fake biometric detection.
- It is useful for cloth pattern recognition for visually impaired blind people.
- It helps in speaker diarylation.
- We can recognize particular objects from different angles.

- **Disadvantages:**
- The syntactic pattern recognition approach is complex to implement and it is a very slow process.
- Sometimes to get better accuracy, a larger dataset is required.
- It cannot explain why a particular object is recognized.
  Example: my face vs my friend’s face.

Block Diagram Explanation:-

1. The sensor converts images or sounds or other physical input into signal data.
2. Segmentation is the part which enables segregating sensed object from the background noise.
3. The feature extractor measures object properties and select those features that are important for classification.
4. The classifier uses these features to assign the sensed object to a category.
5. Finally, the post-processor can take account of other considerations such as the cost of the error to decide appropriate action.



Applications of Pattern Recognition

- **Image processing, segmentation, and analysis**: Pattern recognition is used to give human recognition intelligence to machines that are required in image processing.
- **Computer vision**: Pattern recognition is used to extract meaningful features from given image/video samples and is used in computer vision for various applications like biological and biomedical imaging.
- **Seismic analysis**: The pattern recognition approach is used for the discovery, imaging, and interpretation of temporal patterns in seismic array recordings. Statistical pattern recognition is implemented and used in different types of seismic analysis models.
- **Radar signal classification/analysis**: Pattern recognition and signal processing methods are used in various applications of radar signal classifications like AP mine detection and identification.
- **Speech recognition**: The greatest success in speech recognition has been obtained using pattern recognition paradigms. It is used in various algorithms of speech recognition which tries to avoid the problems of using a phoneme level of description and treats larger units such as words as pattern
- **Fingerprint identification**: Fingerprint recognition technology is a dominant technology in the biometric market. A number of recognition methods have been used to perform fingerprint matching out of which pattern recognition approaches are widely used.

- Imagine we have a dataset containing information about apples and oranges. The features of each fruit are its color (red or yellow) and its shape (round or oval). We can represent each fruit using a list of strings, e.g. [‘red’, ’round’] for a red, round fruit.
- Our goal is to write a function that can predict whether a given fruit is an apple or an orange. To do this, we will use a simple pattern recognition algorithm called k-nearest neighbors (k-NN).
- Here is the function in Python:

```
\from collections import Counter

def predict(fruit):
    # Count the number of apples and oranges in the training data
    num_apples = sum([1 for f in training_data if f[-1] == 'apple'])
    num_oranges = sum([1 for f in training_data if f[-1] == 'orange'])

    # Find the k nearest neighbors of the fruit
    nearest_neighbors = find_nearest_neighbors(fruit, training_data, k=5)

    # Count the number of apples and oranges among the nearest neighbors
    num_apples_nn = sum([1 for nn in nearest_neighbors if nn[-1] == 'apple'])
    num_oranges_nn = sum([1 for nn in nearest_neighbors if nn[-1] == 'orange'])
```

```python
    # Predict the label of the fruit based on the majority class among the nearest neighbors
    if num_apples_nn > num_oranges_nn:
        return 'apple'
    else:
        return 'orange'

# Create a training dataset
training_data = [
    ['red', 'round', 'apple'],
    ['red', 'oval', 'apple'],
    ['yellow', 'round', 'orange'],
    ['yellow', 'oval', 'orange']
]

# Create a test fruit
test_fruit = ['red', 'round']

# Predict the label of the test fruit
prediction = predict(test_fruit)
print(prediction)
```

Output:

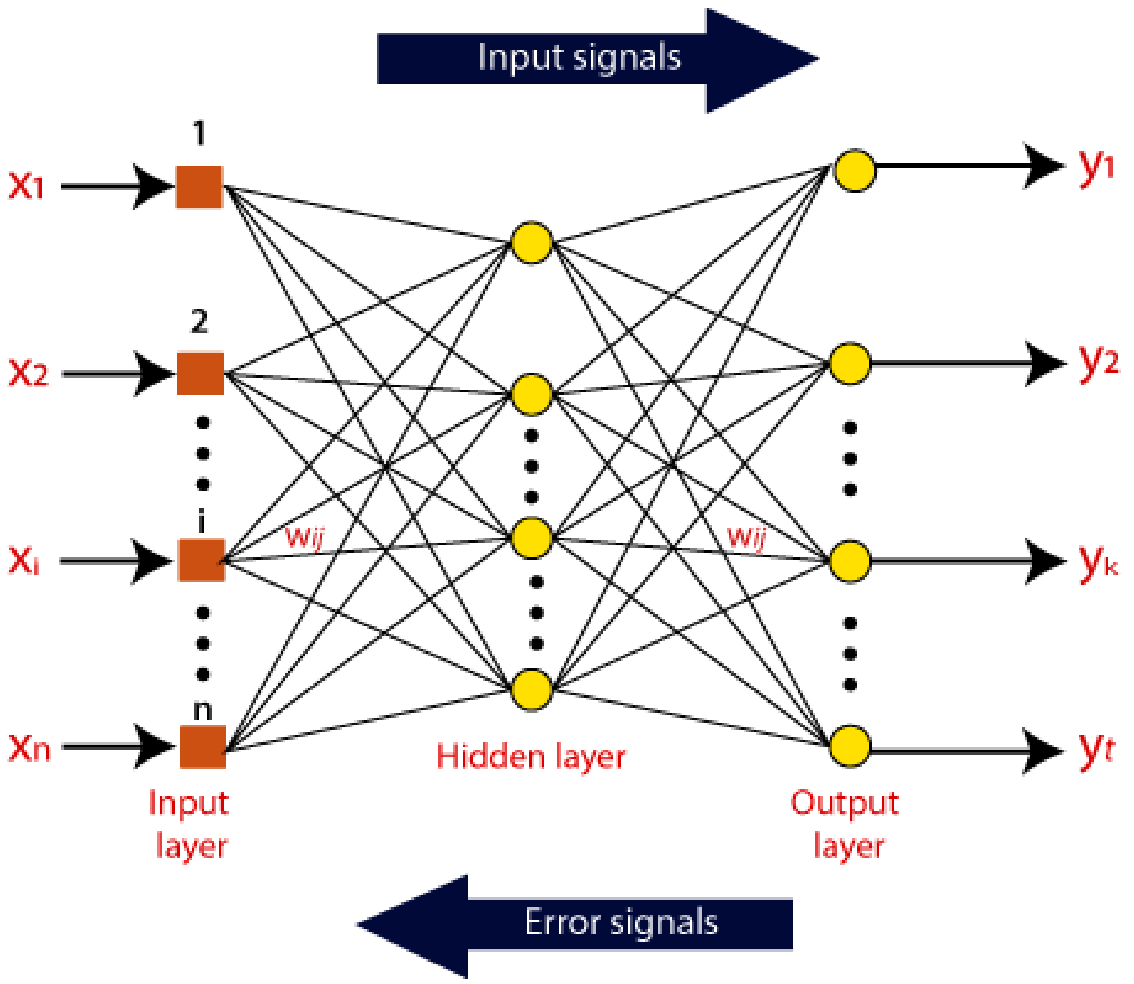
'apple'

What is Neural Network?

- A neural network is a computational model inspired by the way biological neural networks, such as the human brain, work. It consists of interconnected processing units, or "neurons," organized into layers. Each neuron processes information and passes it on to the next layer. Neural networks are designed to recognize patterns, solve complex problems, and make decisions by learning from data.
- Here are the key components and concepts related to neural networks:
  1. **Neurons**: Neurons are the basic processing units that receive inputs, apply weights to them, and produce an output. The output is often transformed using an activation function.
  2. **Layers**: Neurons are organized into layers within a neural network. Common layer types include input layers, hidden layers, and output layers.
  3. **Weights and Biases**: The connections between neurons have associated weights and biases, which are adjusted during training to learn from the data. These weights determine the strength of the connections between neurons.
  4. **Activation Functions**: Activation functions introduce non-linearity to the model. Common activation functions include sigmoid, ReLU (Rectified Linear Unit), and tanh (hyperbolic tangent).
  5. **Feedforward and Backpropagation**: In a feedforward neural network, information flows from the input layer to the output layer. Backpropagation is used to calculate and adjust the gradients of the loss function with respect to the network's weights and biases during training.

Working of ANN

- Artificial Neural Network can be best represented as a weighted directed graph, where the artificial neurons form the nodes. The association between the neurons outputs and neuron inputs can be viewed as the directed edges with weights.
- The Artificial Neural Network receives the input signal from the external source in the form of a pattern and image in the form of a vector. These inputs are then mathematically assigned by the notations x(n) for every n number of inputs.
- Afterward, each of the input is multiplied by its corresponding weights. In general terms, these weights normally represent the strength of the interconnection between neurons inside the artificial neural network. All the weighted inputs are summarized inside the computing unit.
- If the weighted sum is equal to zero, then bias is added to make the output non-zero or something else to scale up to the system's response. Bias has the same input, and weight equals to 1. Here the total of weighted inputs can be in the range of 0 to positive infinity. Here, to keep the response in the limits of the desired value, a certain maximum value is benchmarked, and the total of weighted inputs is passed through the activation function.



Explain how can a neural network be applied in digital image processing with the help of a simple perceptron

- Neural networks, particularly simple perceptron's, can be applied in digital image processing to perform tasks like image classification, edge detection, and noise reduction. Here's a detailed explanation of how a simple perceptron can be used in digital image processing:

Understanding the Simple Perceptron

- A simple perceptron is the most basic type of artificial neural network. It consists of a single layer of neurons (also called nodes) that are fully connected to the input layer. Each neuron performs a weighted sum of its inputs, applies an activation function, and produces an output.

Steps to Apply a Simple Perceptron in Digital Image Processing

1. **Image Representation**:
    1. Digital images are represented as matrices of pixel values. For a grayscale image, each pixel value ranges from 0 (black) to 255 (white).
    2. For simplicity, consider a 3x3 grayscale image, which can be represented as a 9-element vector.

1. **Input Layer**:
    1. The input layer of the perceptron will have one neuron for each pixel in the image. For the 3x3 image, there will be 9 input neurons.
2. **Weights and Bias**:
    1. Each input neuron is connected to an output neuron via a weight. Additionally, a bias term is added to the weighted sum to shift the activation function.
    2. Initialize weights and biases, usually with small random values.

4. **Weighted Sum**:

- Each input pixel value is multiplied by its corresponding weight, and the results are summed up along with the bias term. Mathematically, this is represented as:

$$z = \sum_{i=1}^{n} w_i \cdot x_i + b$$

where x_i are the pixel values, w_i are the weights, b is the bias, and z is the weighted sum.

5. **Activation Function**:

- The weighted sum z is passed through an activation function. Common activation functions include the step function, sigmoid, and ReLU (Rectified Linear Unit).
- For a simple perceptron, a step function can be used to produce a binary output:

$$\text{output} = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{if } z < 0 \end{cases}$$

6. Output:

- The output of the perceptron can be interpreted as a class label in image classification tasks. For example, a perceptron can be trained to distinguish between images of cats and dogs.

Example: Edge Detection

Let's consider a simple edge detection example using a 3x3 image and a perceptron.

1. Input Image:

```
Copy code

1 1 1
0 0 0
1 1 1
```

Weight Initialization: Initialize weights to detect horizontal edges, for instance

```css
w = [1, 1, 1, -1, -1, -1, 1, 1, 1]
```

3. **Bias**:

- Set the bias to a small value, e.g., $b = 0$.

4. **Weighted Sum Calculation**:

- Flatten the 3x3 image to a 9-element vector and compute the weighted sum.

$$z = (1 \cdot 1) + (1 \cdot 1) + (1 \cdot 1) + (0 \cdot -1) + (0 \cdot -1) + (0 \cdot -1) + (1 \cdot 1) + (1 \cdot 1) +$$

$$z = 1 + 1 + 1 + 0 + 0 + 0 + 1 + 1 + 1 = 6$$

5. **Activation Function:**

- Apply the step function:

$$\text{output} = \begin{cases} 1 & \text{if } 6 \geq 0 \\ 0 & \text{if } 6 < 0 \end{cases}$$

- The output is 1, indicating a detected edge.

Conclusion

- A simple perceptron can be effectively used in digital image processing tasks such as edge detection. By carefully selecting and adjusting the weights and biases, a perceptron can learn to recognize patterns and features in images. For more complex tasks, multi-layer perceptron's (MLPs) or deep neural networks (DNNs) are used, as they can capture more intricate patterns and relationships in the data.

- Explain how can a neural network be applied in digital image processing with the help of a simple perceptron
- Discuss Neural Network based image recognition system with the help of a simple perception.
- What are components of a pattern recognition system? Explain the relevant diagrams.
- Write short notes on:(3 + 3)

  a) Pattern Recognition and its applications

  b) Neural Network

- What is pattern and pattern recognition? Explain the steps and application areas of pattern recognition system.
- What is pattern recognition? Explain the applications of neural networks in pattern recognition.