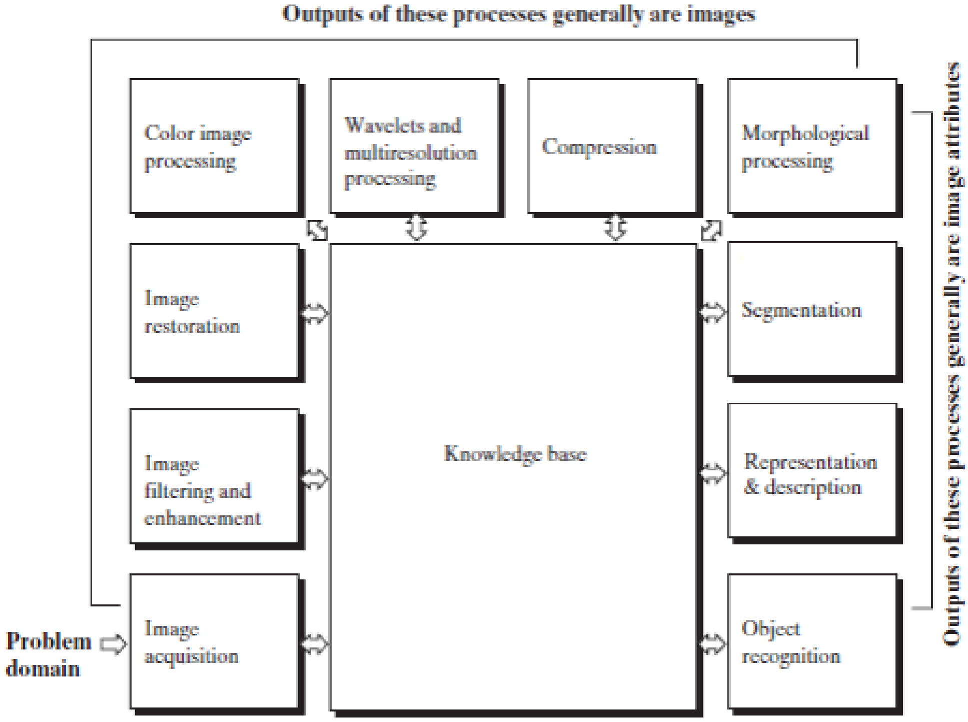
21CSE251T Digital Image Processing

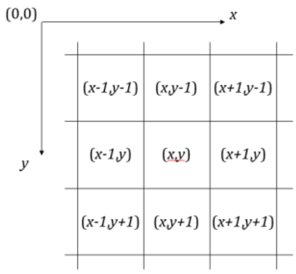
PART – B (5 X 8 = 40 marks)

21 a **Describe the fundamental steps used for processing digital image with block diagram**



21 b **Explain the following relationship between pixels**
(i) Connectivity
(ii) Distance measures

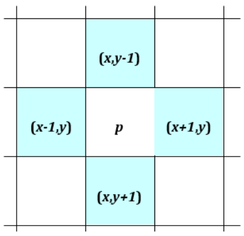
Basic Relationship of Pixels



Conventional indexing method

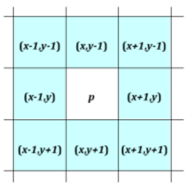
(I) Connectivity

- Neighborhood relation is used to tell adjacent pixels. It is useful for analyzing regions
- 4-neighborhood relation considers only vertical and horizontal neighbors.



4-neighbors of *p*

$$N_4(p) = \left\{ \begin{array}{c} (x-1,y) \\ (x+1,y) \\ (x,y-1) \\ (x,y+1) \end{array} \right\}$$



8-neighbors of *p*

$$N_8(p) = \left\{ \begin{array}{c} (x-1,y-1) \\ (x,y-1) \\ (x+1,y-1) \\ (x-1,y) \\ (x+1,y) \\ (x-1,y+1) \\ (x,y+1) \\ (x+1,y+1) \end{array} \right\}$$

- Two pixels are connected if:
- They are neighbors (i.e. adjacent in some sense -- e.g. N4(p), N8(p), …)

(ii)Distance measures

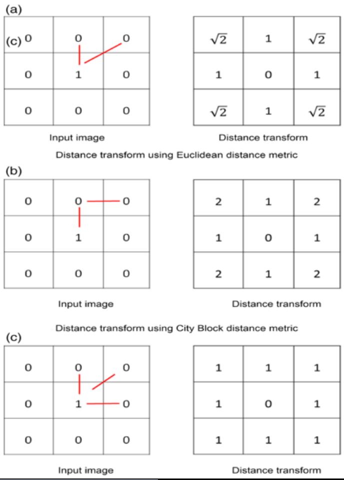
For pixel p, q, and z with coordinates (x,y), (s,t) and (u,v), D is a ***distance function*** or ***metric*** if

- w $D(p,q)^{3} = 0$ ($D(p,q) = 0$ if and only if $p = q$)
- w $D(p,q) = D(q,p)$
- w $D(p,z) \in D(p,q) + D(q,z)$
- The distance transform provides a metric or measure of the separation of points in the image.

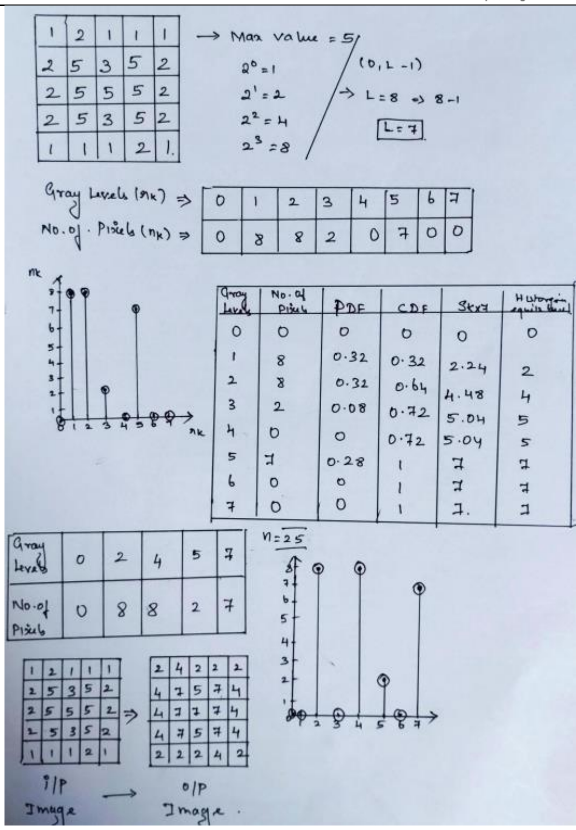
Euclidean Distance $$D_e(p,q) = \sqrt{(x-s)^2 + (y-t)^2}$$

City Block Distance $$D_4(p,q) = |x-s| + |y-t|$$

Chessboard Distance $$D_8(p,q) = \max(|x-s|, |y-t|)$$

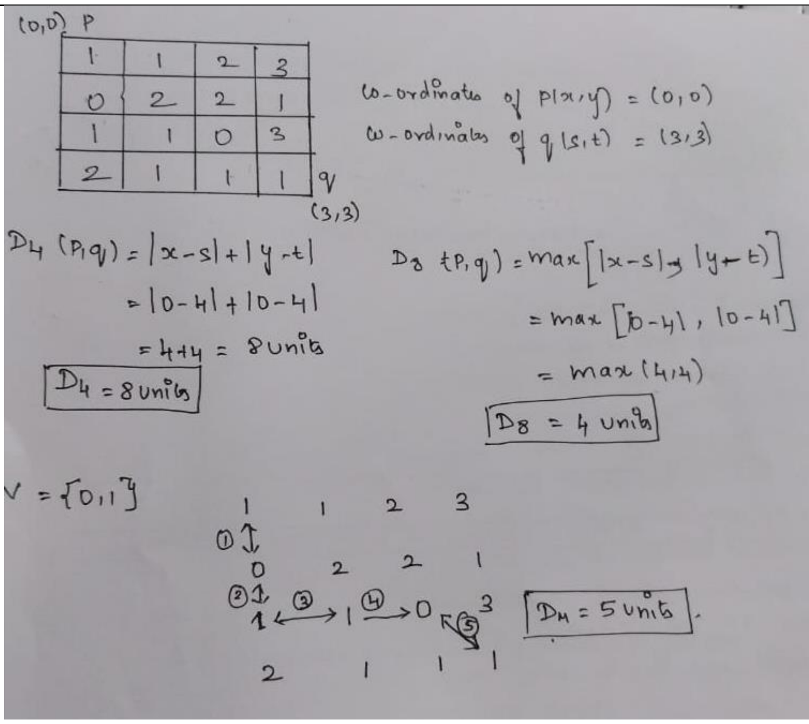


22 A



22 b

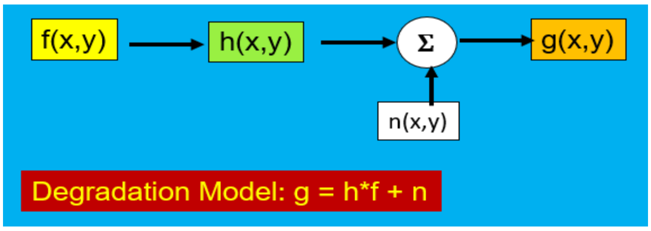
D4= 6 units
D8= 3 units
V= {0,1}= 5 units



23 A

Explain the concepts of degradation model.

- The objective of restoration is to get an estimate f^(x,y) of the original image from g(x,y), the noise term η(x,y) and the degradation function H
- The closeness of the restored image to the original image mainly depends on H & η
- The noise η is unknown, only its statistical property can be learnt



- If H is a linear and position invariant process, then the degraded image in the spatial domain is given by

  $$g(x,y) = h(x,y) * f(x,y) + \eta(x,y)$$

- Where h(x,y) is the spatial representation of the degradation function
- In the frequency domain

  $$G(u,v) = H(u,v)\,F(u,v) + N(u,v)$$

23 B

Derive a wiener filter for image restoration and explain its advantages over the inverse filter

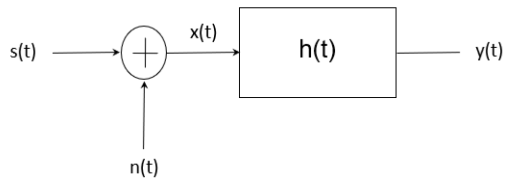
The Wiener filter is a statistical approach that minimizes the mean square error (MSE) between the estimated image and the true image.

- Here images & noise are taken as random processes
- The aim is to find an estimate f^ of the uncorrupted image f such that the mean square error between them is minimum
- This error is given as

$$e^2 = E\{(f - \hat{f})^2\}$$

Where E{.} is the expected value of the argument

Wiener Estimator



Mean Square Error = $\varepsilon\{e^2(t)\}^2 = \int_{-\infty}^{\infty} e^2(t)dt$

$$e(t) = s(t) - y(t)$$

Given the degraded image g, the Wiener filter is an optimal filter h_{win} such that $E\{|| f - h_{win}g||^2\}$ is minimized.
Assume that f and η are uncorrelated zero mean stationary 2D random sequences with known power spectrum S_f *and* S_n. Thus,

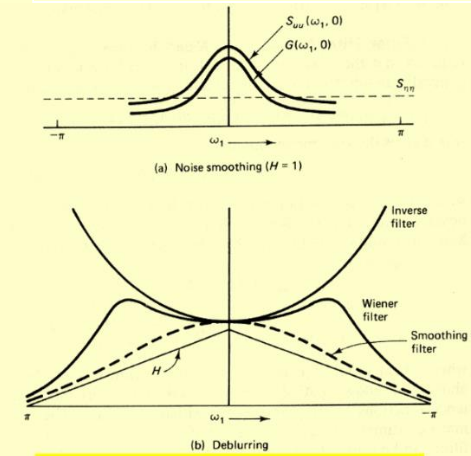
$$E\left\{\|F(u,v)\|^2\right\} = S_f(u,v)$$
$$E\left\{\|N(u,v)\|^2\right\} = S_n(u,v)$$
$$E\left\{F(u,v)N^*(u,v)\right\}$$

- $= E\left\{F^*(u,v)N(u,v)\right\} = 0$

$$H_{win}(u,v) = \left[\frac{H^*(u,v)S_f(u,v)}{S_f(u,v)\left\|\left|H(u,v)\right|^2\right\| + S\eta(u,v)}\right]$$

- It is assumed that the noise & the image are uncorrelated and that they have zero mean
- The gray levels in the estimate are linear function of the levels in the degraded image
- Then the minimum error function in the frequency domain is given as

$$\hat{f}(u,v) = \left[\frac{1}{H(u,v)}\frac{|H(u,v)|^2}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)}\right]G(u,v)$$



Weiner Filter Characteristics

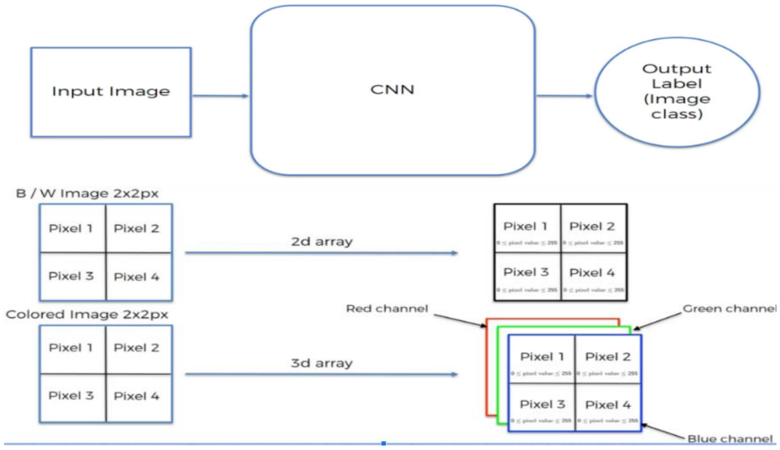
	Advantages of the Wiener Filter over the Inverse Filter • Noise Handling • Stability • Optimal MSE The Wiener filter provides a robust and optimal solution for image restoration by effectively balancing the restoration of the signal and suppression of noise, making it superior to the inverse filter in practical applications.

24 A	**What is the need for image compression? Explain two image compression techniques in detail**
	• Image compression is art and science of reducing amount of data required to represent the image.
	• Data and information are not same
	• Data is a mean to convey the information
	
	Need for image compression
	• To reduce the storage space
	• To reduce the transmission time while web surfing
	• In image compression context, a in the equation for C_R, usually is the number of bits needed to represent an image as 2-D array of intensities
	• Typically, 2-D array of intensities suffer from three types of redundancies
	  • Coding Redundancy
	  • Spatial or temporal redundancy
	  • Irrelevant information
	Image Compression Techniques
	
	Lossless Image Compression Techniques
	
	• Run-Length Encoding (RLE):
	• Huffman Coding:
	• Arithmetic coding
	Lossy Image Compression Techniques
	• Transform Coding:
	• Quantization:
	• Gray-Level Run Length Coding

24 B	**Explain the following features** **(i) Histogram based features** **(ii) Intensity Features** (i) **Histogram features** • V is the set of gray-level values used to define adjacency (e.g. V={1} for adjacency of pixels of value 1) • *S*: a subset of pixels in an image. • Two pixels *p* and *q* are said to be connected in *S* if there exists a path between them consisting entirely of pixels in *S*. • For any pixel *p* in *S*, the set of pixels that are connected to it in *S* is called a connected component of *S*. • Their gray levels satisfy a specified criterion of similarity (e.g. equality, …)

	• Histogram features are also known as brightness features and are described in terms of the luminance (or intensity) features. • The histogram shows the brightness distribution found in the object. • The first order histogram of an image can be approximated using the below equation $$p(b) = \frac{N(b)}{m}$$ Here, N(b) is the number of pixels having an amplitude of r_b L is the total grey levels and m is the total number of pixels in the image. **First order moments** First order moments are calculated based on individual pixel values. Some of the features that are helpful include the following. 1. Mean 2. Standard deviation 3. Skewness 4. Kurtosis 5. Energy 6. Entropy **Second Order Moments.** • These are based on joint probability distribution of pixels I and J. The histogram of histogram of the second order distribution is given by $$\boxed{p(a,b) \approx \frac{N(a,b)}{m}}$$ Where, m is the total number of pixels. These measures are useful in measuring the texture of images and includes Autocorrelation, Covariance, Inertia, Absolute value, inverse difference, energy and entropy.
25 A	**Describe CNN and demonstrate how it detects objects** Convolutional neural networks are deep artificial neural networks that are used primarily to classify images, cluster them by similarity (photo search), and perform object recognition within scenes. By use of CNNs it can identify faces, individuals, any signs, tumors and many other aspects of visual data. **Importance of CNN:** • Convolutional Neural Networks are complex feed forward neural networks. CNNs are used for image classification and recognition because of its high accuracy. • It was proposed by computer scientist Yann LeCun in the late 90s, when he was inspired from the human visual perception of recognizing things.

- The CNN follows a hierarchical model which works on building a network, like a funnel, and finally gives out a fully-connected layer where all the neurons are connected to each other and the output is processed.

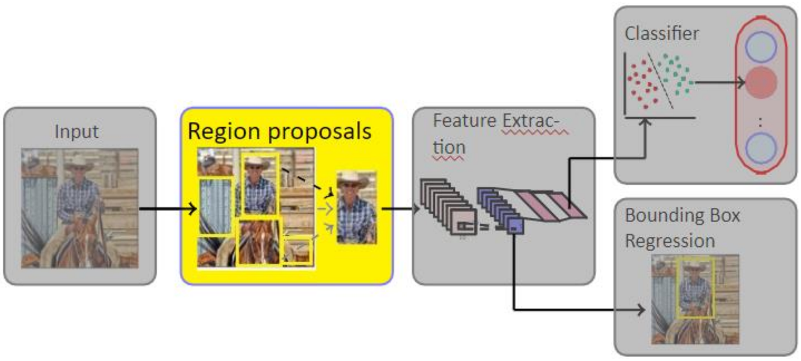


Layers:

- Convolutional layers
  - Feature Map or filter
  - Shared weights
- Subsampling or Max pooling
- Full connected layer (classification)

Region-based Convolutional Neural Networks (R-CNN):

R-CNN: It generates region proposals using selective search, warps each proposal to a fixed size, and then classifies each region using a CNN.



25 B

What is binary classifier and how does SVM work in image classification? Explain it in a neat block diagram

Binary classifier is a type of machine learning model that categorizes images or parts of images into one of two categories: typically, a positive class (e.g., "contains object of interest") and a negative class (e.g., "does not contain object of interest").

Image classification using SVM

- A support vector machine (SVM) is a type of supervised learning algorithm used in machine learning to solve classification and regression tasks;

- A new classification method for both linear and nonlinear data. SVM is applicable for the data that are linearly separable. In Non-linear data kernel functions are used. SVM finds this hyperplane using support vectors ("essential" training tuples) and margins (defined by the support vectors).

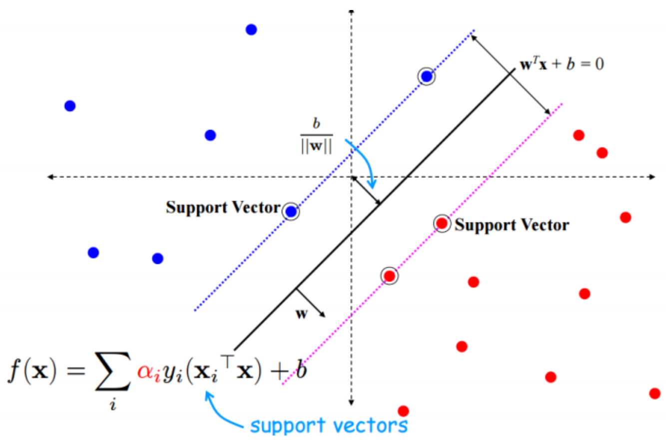
- SVMs can be used for a variety of tasks, such as text classification, image classification, spam detection, handwriting identification, gene expression analysis, face detection, and anomaly detection

Properties of SVM

- 1. SVMs construct a maximum margin separator—a decision boundary with the largest possible distance to example points

- 2. SVMs create a linear separating hyperplane, but they have the ability to embed the data into a higher-dimensional space, using the so-called kernel trick. Often, data that are not linearly separable in the original input space are easily separable in the higher dimensional space.

- 3. SVMs are a nonparametric method—they retain training examples and potentially need to store them all. SVMs combine the advantages of nonparametric and parametric models: they have the flexibility to represent complex functions, but they are resistant to overfitting.



PART – C (1 x 15= 15 Marks)

26

7	9	5
4	6	8
2	0	1

Box or mean filter= [7+9+5+4+6+8+2+0+1] / 9 = **4.6 or 5**

Weighted Average filter = [7+9+5+4+12+8+2+0+1] / 9= **5.3 or 5**

Median filter = *0,1,2,4,***5***,6,7,8,9* = **5**

Min filter = 0,1,2,4,**5**,6,7,8,9 = **0 (minimum value)**

Max filter = 0,1,2,4,**5**,6,7,8,9 = **9 (Maximum value)**

27 **Design a system for detecting drowsiness using image processing techniques , suggest a suitable algorithm for each step.**

	Designing a system for detecting drowsiness using image processing involves several steps, from capturing the image to making a drowsiness decision. **Image Acquisition:** Capture video frames using a camera. **Face Detection:** Detect faces in the frames using Haar Cascades or MTCNN. **Eye Detection:** Detect eyes within the face region using facial landmark detection. **Eye State Classification:** Classify the eye state (open or closed) using a CNN. **Drowsiness Detection Logic:** Use the PERCLOS method to determine drowsiness based on the percentage of time eyes are closed. This system will help in real-time drowsiness detection by processing video frames, detecting and classifying eye states, and applying logical rules to infer drowsiness.