

Image processing basics

Dr. Sagarika B

Agenda : Session 4

- Image representations
- Linear filtering
- Organize image data for computer vision problems

What is an image ?

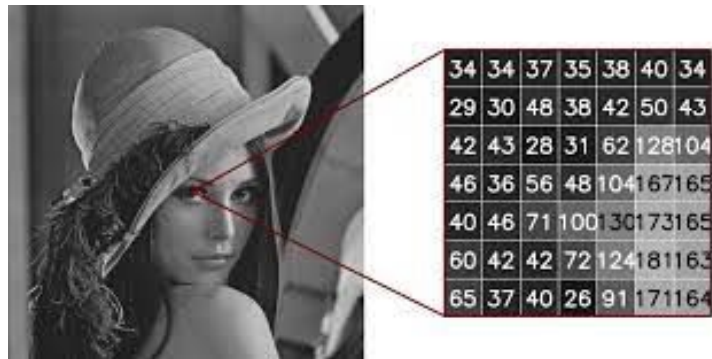
Pixel grid

How they are stored in memory?

The typical format for storing pixels is by byte or eight bits. 8 bits can store 2^8 amount of information. Grayscale images have 1 channel that goes from a scale from 0 to 255 where 0 is black and 255 is white. Colored images typically have 3 channels and stored as 3 bytes: red, green, blue (RGB) values each ranging from 0 to 255 depending on intensity.

Image representations

- **Grayscale representation:** Images are represented using a single channel where each pixel contains a grayscale value ranging from 0 (black) to 255 (white). This representation is commonly used for tasks that do not require color information.



Color representation:

RGB color Space : For color images, the most common representation is the RGB (Red, Green, Blue) format. Each pixel is represented by three color channels (R, G, and B), with each channel containing an intensity value ranging from 0 to 255.



Other color spaces and their usage :

Monitors: RGB

Printers: CMY

Human perception: HSI [hue (the “inherent/pure” color – red, orange, purple, etc.), saturation (the amount of white mixed in the color, i.e. pink versus magenta), intensity (the amount of black mixed in the color, i.e. dark red versus bright red).]

Efficient compression and transmission: YCbCr.

The YCbCr color space is a similarly decorrelated color space with Y being the luminance channel similar to the V in HSV. The luminance channel (Y) carries most information from the point of view of human perception, and the human eye is less sensitive to changes in chrominance.

Image Basics

Image size

Image size refers to the dimensions of an image, typically measured in pixels, as well as the file size, measured in kilobytes (KB), megabytes (MB), or gigabytes (GB).

Bit Depth

Bit depth refers to the number of bits used to represent the color of a single pixel. It determines the number of colors an image can display.

- **1-bit:** Black and white images with 2 possible colors (2^1).
- **8-bit:** Grayscale images with 256 possible shades of gray (2^8).
- **24-bit (True Color):** Images with 16.7 million possible colors (2^{24}), with 8 bits for each RGB channel (Red, Green, Blue).

What is Image Resolution?



- ★ Image resolution refers to the amount of detail an image holds. It is typically measured in terms of pixels, with higher resolutions indicating more detail and clarity

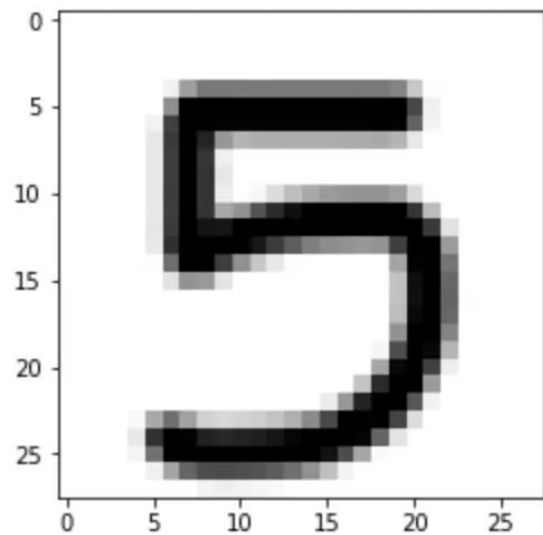
Standard Definition (SD): 640x480 pixels.

High Definition (HD): 1280x720 pixels.

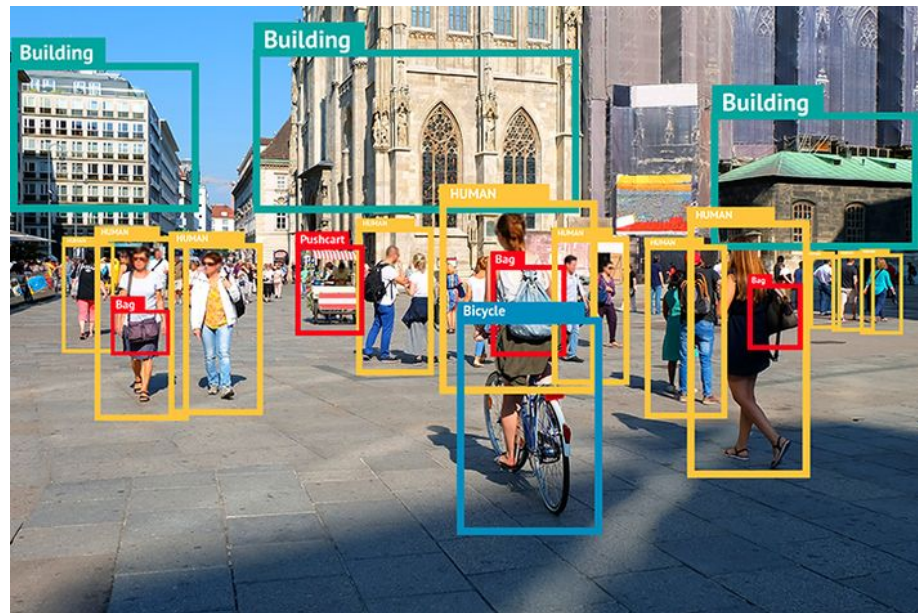
Full HD (FHD): 1920x1080 pixels.

4K Ultra HD (UHD): 3840x2160 pixels.

8K Ultra HD (UHD): 7680x4320 pixels.



VS



Traditional Image Formats



Raster Images

Pixel-based graphics
Resolution dependent
Photos & web graphics

JPG

Web & print
photos and
quick previews

GIF

Animation &
transparency in
limited colors

PNG

Transparency
with millions
of colors

TIFF

High quality
print graphics
and scans

RAW

Unprocessed
data from
digital cameras

PSD

Layered Adobe
Photoshop
design files



Vector Images

Curve-based graphics
Resolution independent
Logos, icons & type

PDF

Print files and
web-based
documents

EPS

Individual
vector design
elements

AI

Original Adobe
Illustrator
design files

SVG

Vector files
for web
publishing

Python Libraries for Image Processing

01

Scikit-Image

02

OpenCV

03

NumPy

04

PIL/pillow

05

Mahotas

What is Image filtering

Most common filters are linear filters and the process of applying a linear filter is called convolution

Why filter –

Enhance images : Denoise, resize, increase contrast, etc.

Extract information from images •

Texture, edges, distinctive points, etc.

Linear filters

General process:

– Linear filtering is the filtering method where the value of output pixel is linear combinations of the neighboring input pixels. It can be done with convolution operation. For example, mean/average filters or Gaussian filter.

- Properties

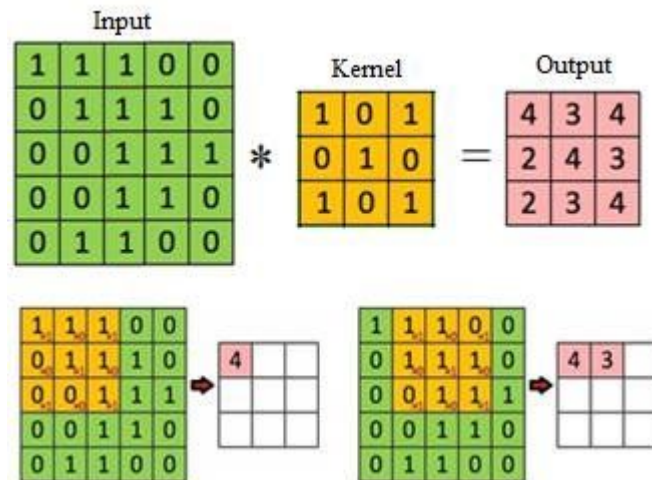
- Output is a linear function of the input

- Output is a shift-invariant function of the input (i.e. shift the input image two pixels to the left, the output is shifted two pixels to the left)

Convolution

In image convolution, involves a kernel, or matrix that is applied over the input image's pixels to generate an output image. The kernel size and values determine the effect the kernel has on the image. The dimensions of the kernel should be smaller or equal to that of the input image's.

The darkest pixels represent the smallest results of the operation between a receptive field and the kernel, on the other hand, the brightest pixels represent the highest values for the operation between a receptive field and the kernel.



Edge detection (using convolution)

- Edges are sudden discontinuities in an image, which can arise from surface normal, surface color, depth, illumination, or other discontinuities.

Edges are important for two main reasons.

- 1) Most semantic and shape information can be deduced from them, so we can perform object recognition and analyze perspectives and geometry of an image.
- 2) They are a more compact representation than pixels.

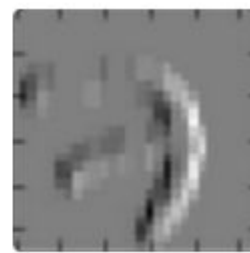
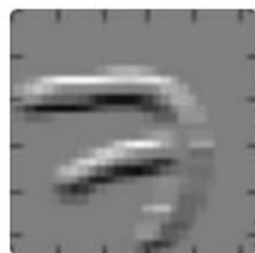
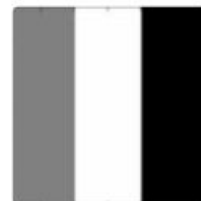
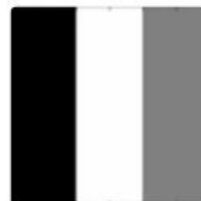


-1	-1	-1
1	1	1
0	0	0

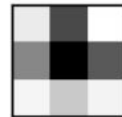
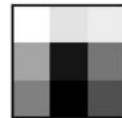
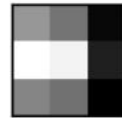
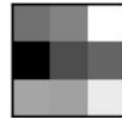
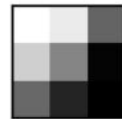
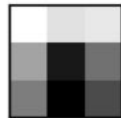
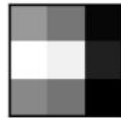
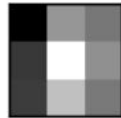
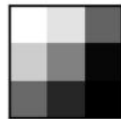
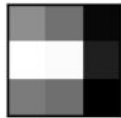
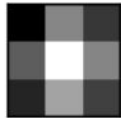
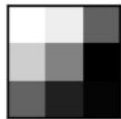
-1	1	0
-1	1	0
-1	1	0

0	0	0
1	1	1
-1	-1	-1

0	1	-1
0	1	-1
0	1	-1



Filters !!!



Other linear filters :

Box Blur: Simple averaging, reduces noise and detail.

Gaussian Blur: Natural blurring effect, minimizes edge distortion.

Sharpening: Enhances edges and details.

Emboss: Creates a 3D-like effect by highlighting edges.

Motion Blur: Simulates camera motion.

Unsharp Masking: Enhances edges by subtracting a blurred image.

Sobel Filter: Enhances edge contrast in specific directions.

Features or patterns !!

These are what we mean by *patterns*.

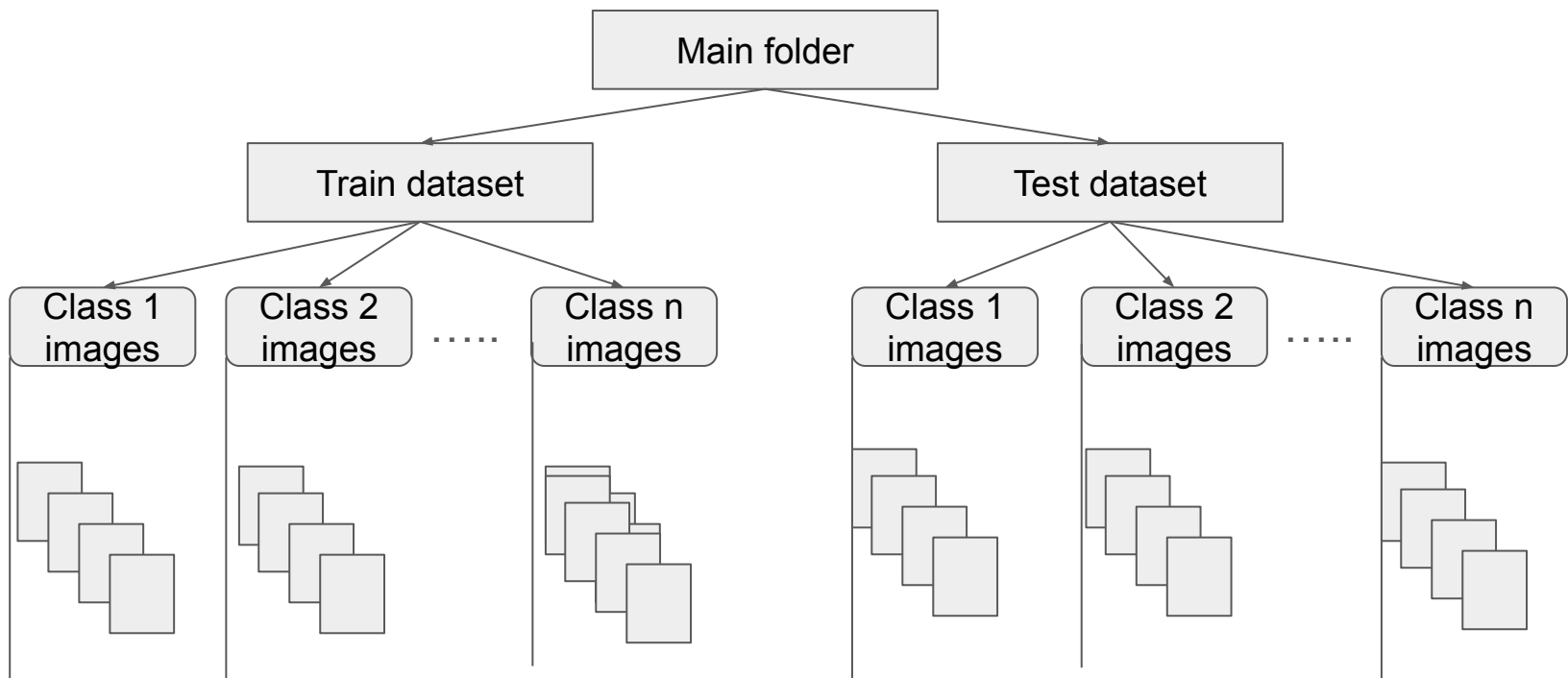
- edges
- shapes
- textures
- curves
- objects
- colors

- Aside from edges, some filters may detect corners. Some may detect circles. Others, squares. Now these simple, and kind of geometric, filters are what we'd see at the start of a convolutional neural network.
- The deeper the network goes, the more sophisticated the filters become. In later layers, rather than edges and simple shapes, our filters may be able to detect specific objects like eyes, ears, hair or fur, feathers, scales, and beaks.
- In even deeper layers, the filters are able to detect even more sophisticated objects like full dogs, cats, lizards, and birds.

Lets visualize the filtering operations

How to Organize
images for computer
vision problems?





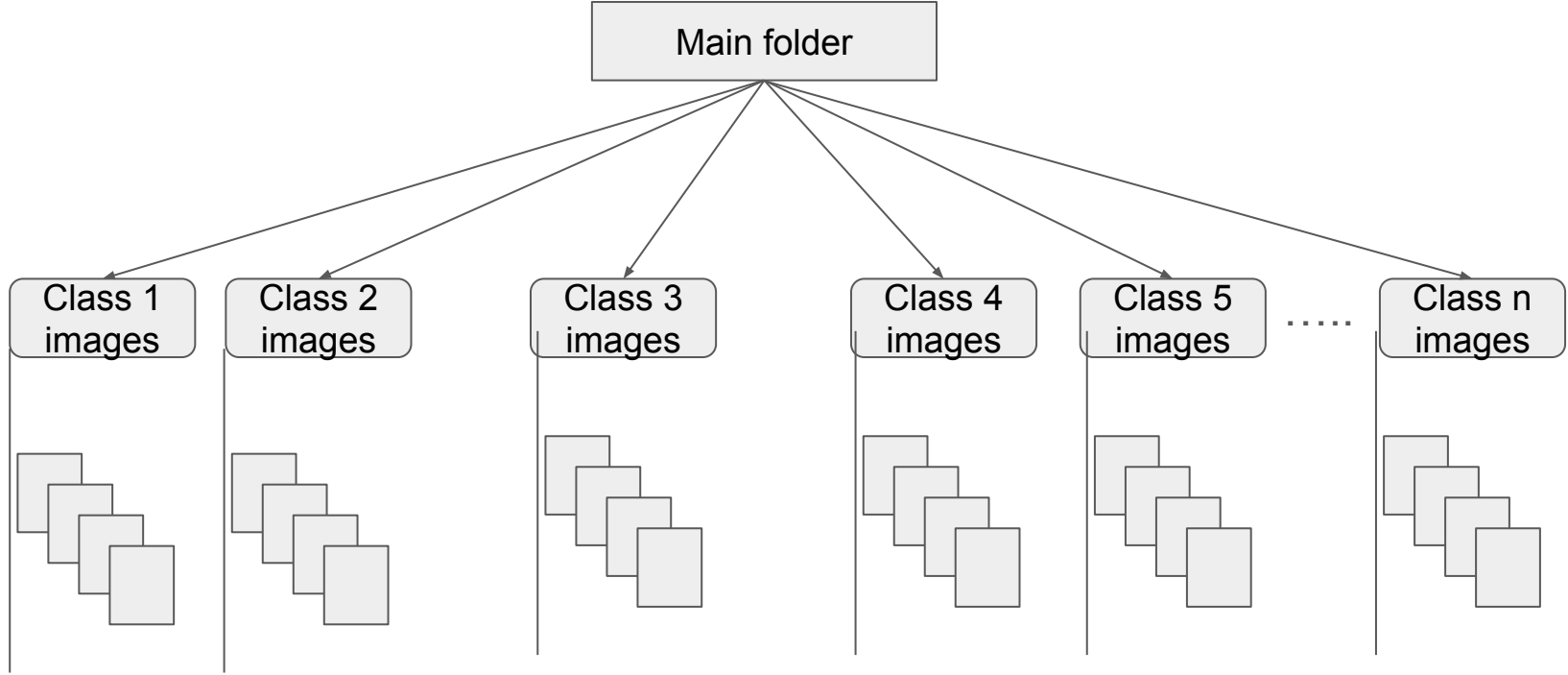
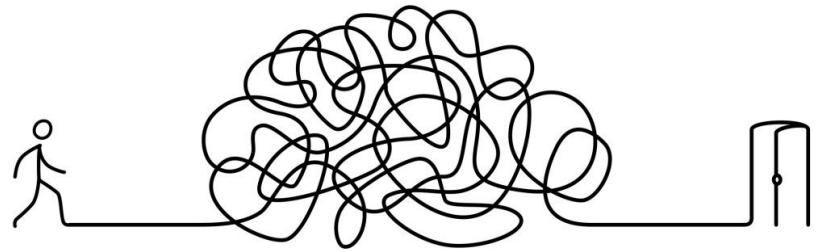


Image info can be stored in CSV file and read through it.

Real time image data challenges !!

- Very large sized image with lot of details
- Input Images of different size
- Image, label mapping might be wrong (Domain expertise might be needed to verify)
- Too many images to process



Good materials

<https://ai.stanford.edu/~syYeung/cvweb/tutorial1.html#:~:text=Image%20filtering%20changes%20the%20range,points%20without%20changing%20the%20colors.>

OpenCV documentation

https://docs.opencv.org/4.x/d2/d96/tutorial_py_table_of_contents_imgproc.html