

COMPUTER VISION

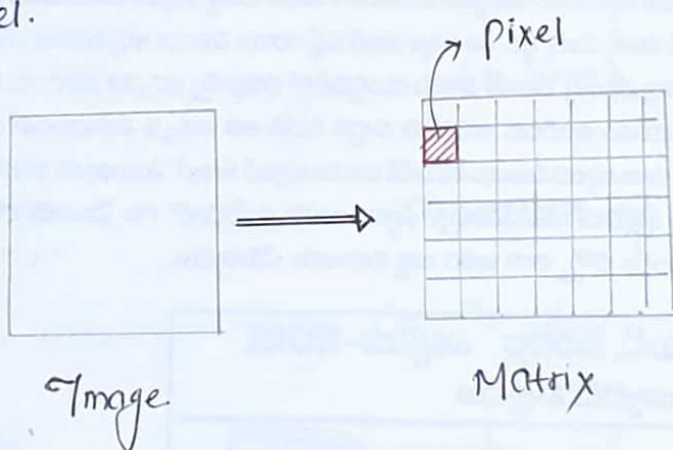
①

Computer vision is a field of artificial Intelligence (AI) that enables Computers and Systems to derive meaningful information from digital images, videos and other visual inputs and take actions to make recommendations based on the information.

Multimedia :- it is the use of Computers to present and combine text, graphics, audio and video with links and tools that let the user navigate, interact, create and communicate.

Image :- Image is nothing but matrices which contain numerical dataset and it is a visual representation of something.

Matrix contains row and column representation, where numerical values are stored, this whole set of values stored in a container, which is generally called as pixel.



Generally, pixels tell us about color saturation, i.e., whether an image is black, white, or grey, primarily.

Pixels :- A pixel is one of the small dots or squares that makes up an image on a computer screen. The more pixels there are, the more the image looks real or accurate, any digital image is made up of pixels and when someone talks about resolution of computer monitor or TV screen, they are referring to the no. of pixels.

Note :- Generally an image is matrix of zeros called as black image. ie entire image is in color of black. if its pixels contain matrix of 255 called as white image, the mean of it represent grey color.

$$\begin{matrix} 0 \rightarrow \text{black color} \\ 255 \rightarrow \text{white color} \end{matrix} \left. \vphantom{\begin{matrix} 0 \\ 255 \end{matrix}} \right\} \frac{0+255}{2} = 127 \rightarrow \text{Grey color.}$$

Binary image consist of combination of black and white color. ie there is no color saturation present on it.

$$\text{Binary Image} = (0, 255)$$

Size of an image :- which is description of height and width of an image in pixels. maximum image size is determined by the mega pixels of a given camera.

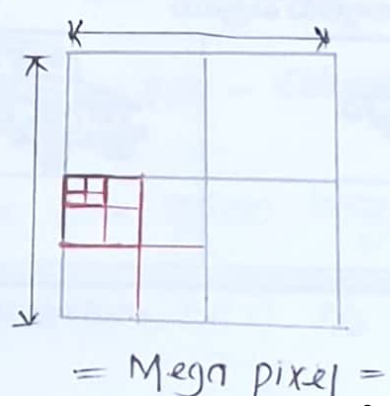
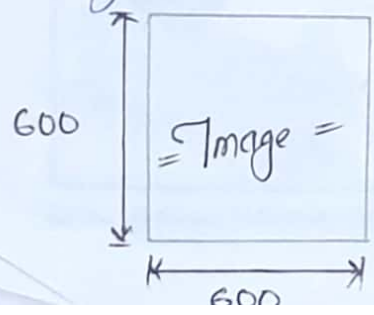


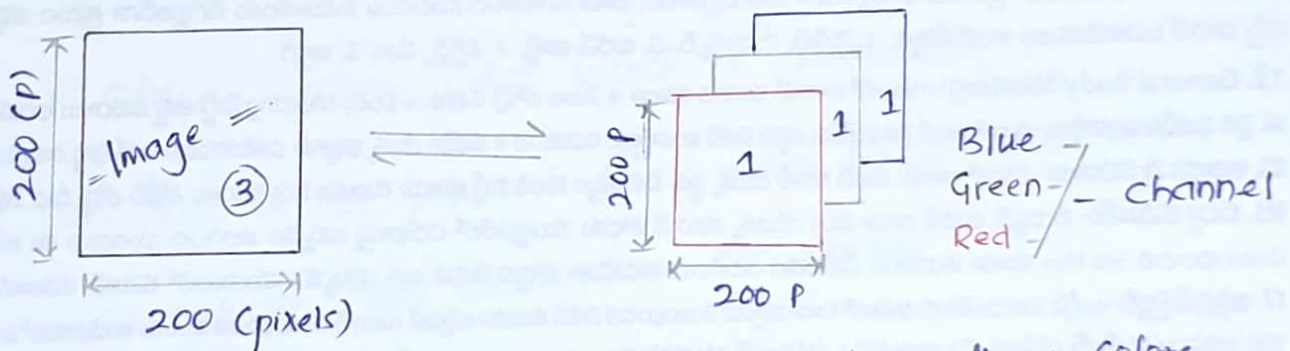
Image Containing 600 Rows and 600 Columns (height and width)

Generally represent size of an image which is equal to number of pixels.

$$[600 \times 600] = \text{no of pixels} \approx \text{Size of an Image}$$

Generally mega pixel tells us more capture. if an image contain more pixels termed as mega pixel and which containing more clarity, the best image size is $3.14 \times 10^6 \Rightarrow 3.4$ mega pixels.

Color :- Color in a computer program are represented by combination of three channels namely as (red, green, blue) which contrasts with the primary color. by combination of certain amount of Red, some amount of Green, some amount of Blue, we get any color.



primary image is Mixer of these three colors

i) Size of an Image = $200 \times 200 \times 3$

Where 200 - Rows (pixels), 200 - Columns (pixels), 3 channels.

ii) if an empty image the size of an image is $200 \times 200 \times 1$

empty means (either black or white) no color saturation.

iii) Generally channels are called as filters, the representation of an color-filter is in form of R.G.B

if an Image is in Red Color then the representation of color is — $(255, 0, 0)$
R G B

Where Red is present, Green absent, Blue absent ie there is none of amount of colors of Green, Blue. only red present that shows red color of an Image.

Similarly Green Color Representation — $(0, 255, 0)$

Blue Color — $(0, 0, 255)$

Black Color — $(0, 0, 0)$ Matrix of zeros.

White Color — $(255, 255, 255)$

Yellow Color — $(255, 255, 0)$

if we observe yellow color which is combination of red and green that's why blue is not present. Some times we get pale colors ie some amount of color combinations are added

if a pixel is nearer to 0 ie $\{1, 2, 3, 4\}$ behaves like black color, if colors nearer to 255 represent white color.

Generally the shape of an Image = $200 \times 200 \times 3$

Shape of Binary Image = $200 \times 200 \times 1$

Bits per pixel (B.p.p) :- B.p.p or Bits per pixel denotes the no of bits per pixel. the no of different colors in an image is depends on the depth of color or bits per pixel.

In mathematically we represent bit in-form of numbers (0,1) two numbers can be represented by one bit.

Combination can be made as- ([00], [10], [01], [11])

$$\text{Bits per pixel} = 2^k$$

where k denotes - b.p.p. if we put $k=1$ we get 2 colors

if $k=2 \Rightarrow 4$ colors combination (ie k - no of colors combination)

$k=3 \Rightarrow 8$ colors are present in the pixel.

Video :- Sequence of an image running very fast, this whole process known as video. Generally we known image as frame we find out the speed in time ie playing of an image per second.

Generally termed as 60 fps ie 60 frames per second. Video is dynamic in nature, this can tell us quality of a video. if we have a low quality video appear that means frame per second considered very low. if we want to improve the quality of a video we have to improve the fps speed. (ie Increase video resolution)

Introduction TO Natural Language Processing :- (NLP) ①

Supervised, Unsupervised Techniques :- Supervised learning algorithms are trained using labeled data i.e. (contain output data) where unsupervised learning algorithms are trained using unlabeled data i.e. (only input data provided to the model)

How to Use NLP :- NLP is broadly defined as the automatic manipulation of natural language, like text and speech, by software, this data comes from in form of image or text by different websites, reviews, etc.

For Suppose if we contain text data of reviews of airlines as shown below -

	Input (Review)	Output (Sentiment)
1	I Like to travel in Saudi Airlines.	Positive
2	I Lost my baggage there, it was very bad experience with me.	Negative
3	I didn't get any observation 😎 (cool)	Neutral

Table of Data represent input as review of customer, Output as Sentiment of that review either positive, negative or neutral, if we observe label contain then we can easily say we have to perform supervised technique in that we have to use "classifier's" algorithms because label containing categorical type of data. From this we perform encoding techniques to convert this data into number for easy

analysis of Sentiments - from review by machine learning algorithm. but our problem is input data (review) this contains text data, we don't have any usage of this data in machine learning to predict some output because we can't understand anything from this, so where we use "NLP" to handle text data, image data, etc.

from this we can say our review 1 containing some positive sentiment, review (2) → Negative, review (3) → nothing i.e. neutral, on top of it if we apply machine learning algorithms we easily find out sentiments of reviews, in order to get any new query point we can say sentiment of that point and analyse sentiments.

finally we can say in order to handle text data we use "NLP" terminology where we convert entire text into mathematical formulation. There are some techniques in "NLP" to convert text data into numerical representation. by usage of it we can easily convert data.

They are —

- 1) BOW → Bag of words
- 2) TF-IDF → Term frequency - Inverse document frequency
- 3) Word 2 Vect → Word vector
- 4) BERT → Bidirectional Encoder Representation from Transformers.