Introduction of Image Processing

Achmad Fauzi Bagus Firmansyah

Direktorat Pengembangan Metodologi Sensus dan Survei, Badan Pusat Statistik

01. What is “Image”?

02. You “SEE” what you care...

03. Tweaking images...

What Will we learn today ?

What is “Image”

IMAGE

Before we dive into the world of computer vision, we need to understand a simple, yet powerful concept — **the image**.

But for a computer, an image is not a picture;
IT'S DATA, represented by matrix of numerical value.

At glance, we se number 7

IMAGE

Representation of top left in Matrix Format:

```
[[  3   0   7   0   2   4   0   0   2   5   3]
 [  2   0   7   0   3   0  10   8   0   0   0]
 [  4   0   0   3  10   0   0   0   0   7   1]
 [  0   5   4   3   0   0  15  18   0   0   0]
 [  0   6   0   0   5   4   0   0   2   2   0]
 [  1   2   0   0  11  11   0  20   7   0   0]
 [  0   6  12   0   0   0  18 192 180  98  29]
 [  2   0   0   1  11   0   0 196 255 245 236]
 [  0   2   1   0   0   0  10  19 183 242 244]
 [  0   1   2   0   0   0   4   9   0  29  82]
 [  0   0   3   2   0   0   0   0  10   0   0]]
```

Expert Definition

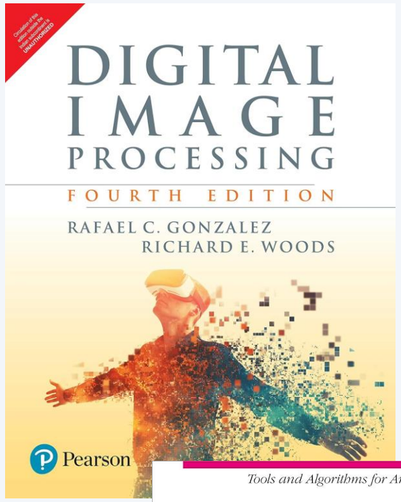
In the context of computer vision, an image is:

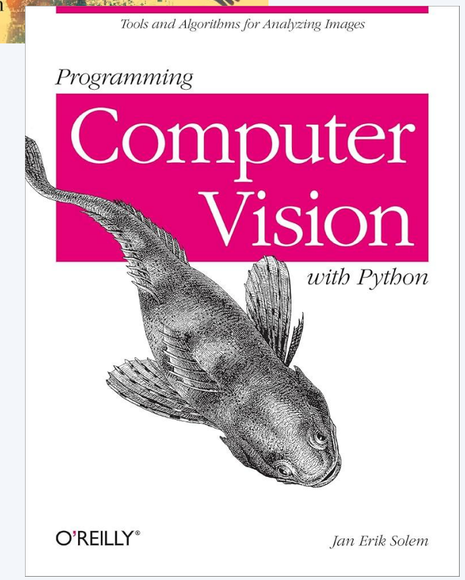
- *“A two-dimensional function $f(x,y)$, where x and y are spatial coordinates, and the amplitude of f at any pair (x,y) is called the intensity or gray level or pixels of the image at that point.”*

  — Gonzalez & Woods, "Digital Image Processing", 4th Edition, Pearson, 2018

- *“An image is a matrix of pixel values that can be processed to extract features, objects, or other interpretable data.”*

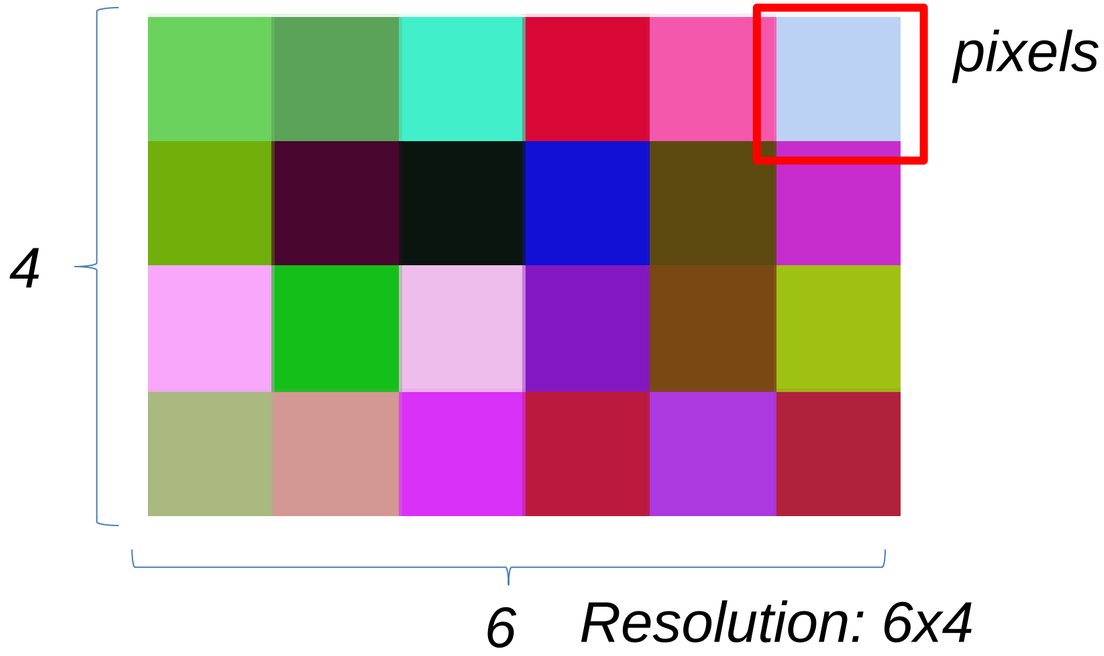
  — Solem, "Programming Computer Vision with Python", O’Reilly, 2012



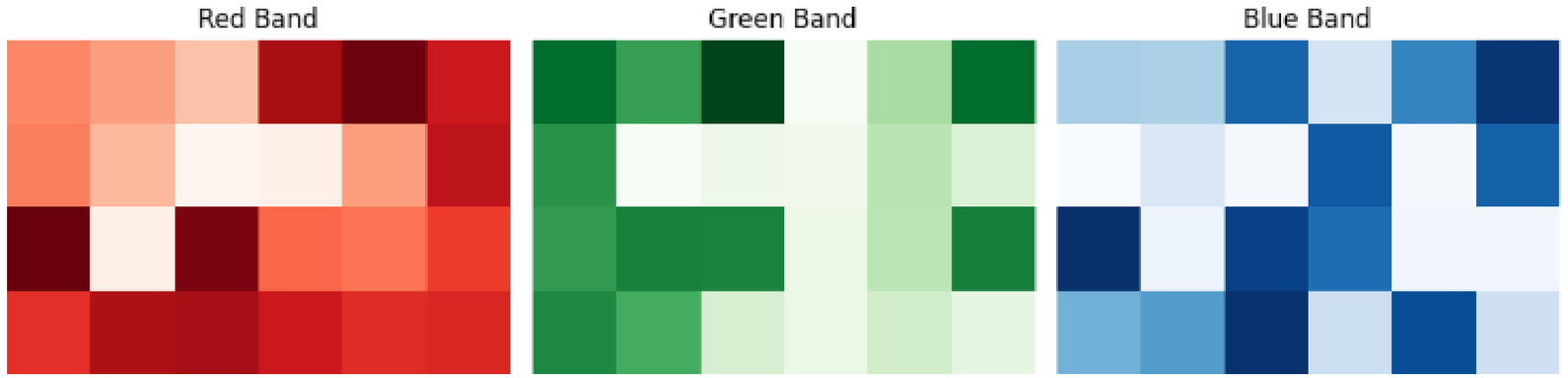


Pixels, Resolution, Band

- **Pixels** is smallest unit of picture that compose an image.
- **Resolution** is size of the image, represented with *length* x *wide*.
- **Band** is layer of information of image, sometime called “channel”, eg: RGB.



Band (RGB)



You “CARE” What you “SEE”

“SEEING”

In human terms, seeing means more than receiving light — it implies **understanding**. Our brains interpret the light signals captured by our eyes **to recognize objects, motion, depth, and patterns**.

In computer vision, **"seeing"** is the **process of converting raw image data into high-level understanding** — for example, detecting a face, identifying a car, or interpreting a handwritten digit.

Try to look this image in 20 seconds

CARE-01

Which of the following categories best describes the picture?

- ***Group picture***
- ***Scenery***
- ***Environment***
- ***Building***
- ***People***
- ***Garden***

CARE-01

Which of the following categories best describes the picture?

- ***Group picture***
- ***Scenery***
- ***Environment***
- ***Building***
- ***People***
- ***Garden***

Here, we attempt to "classify" the images based on our "knowledges".

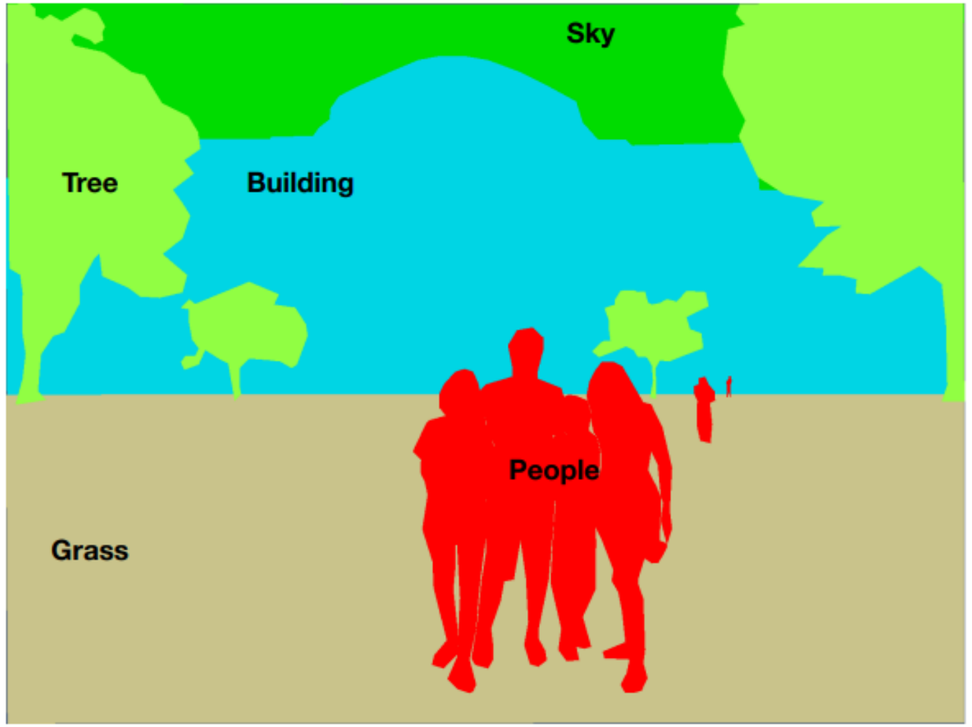
CARE-02

What objects are shown in the picture?

- ***Building***
- ***People***
- ***Sky***
- ***Tree***
- ***Grass***

CARE-02

What objects are shown in the picture?



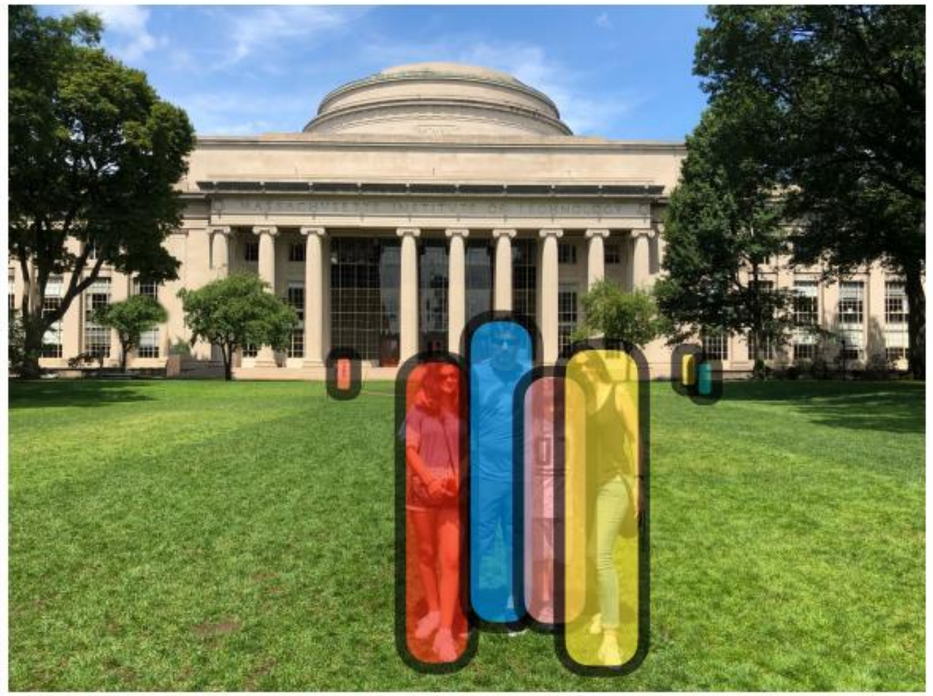
Here, we attempt to "segment" the "object" based on our "knowledges".

CARE-03

How many people are shown ?

CARE-03

How many people are shown ?



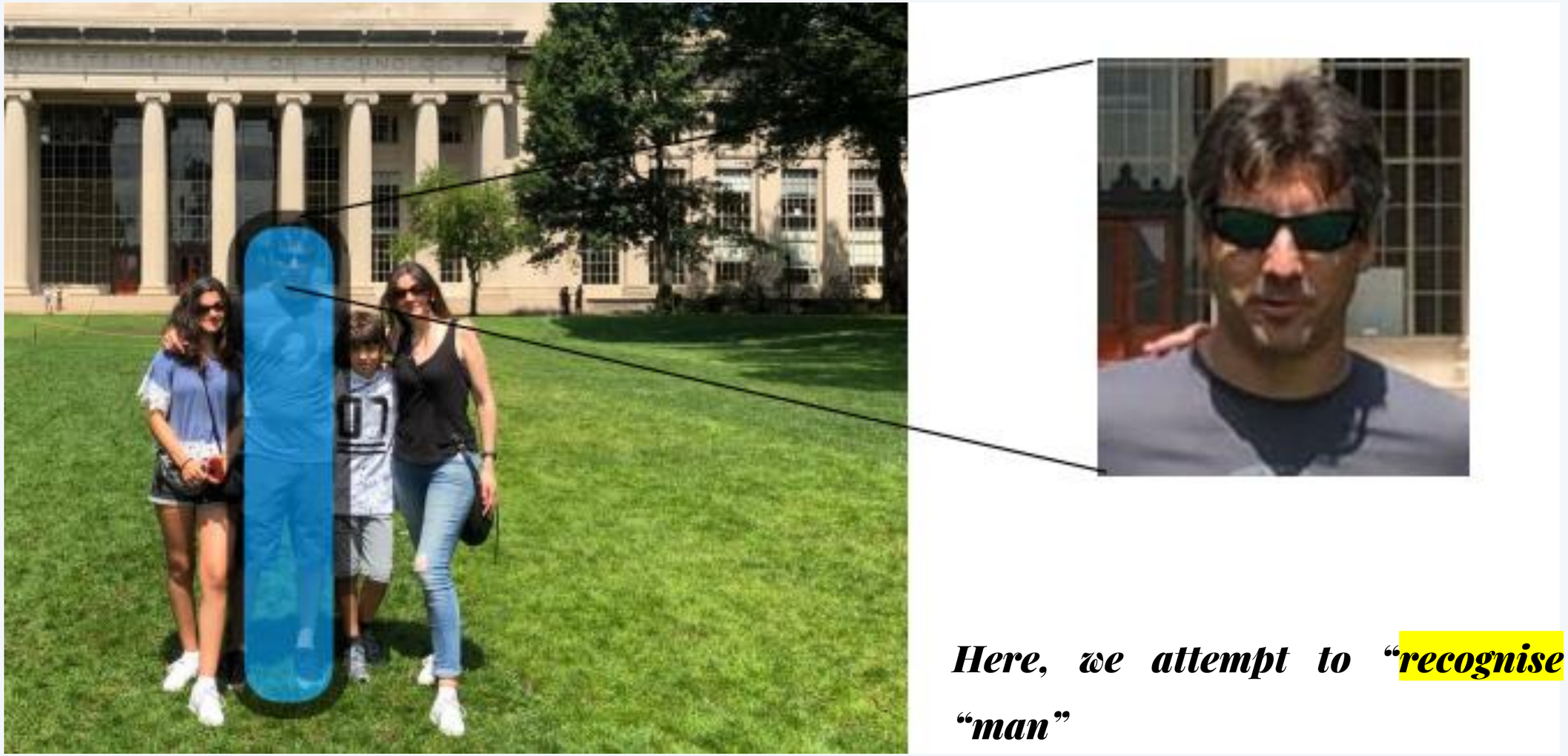
Here, we attempt to "detect" the "people" and "count"

CARE-04

Do you know him?

CARE-04

Do you know him?



Here, we attempt to "recognise" the "man"

CARE-04

Could you tell me about this picture?

CARE-04

Could you tell me about this picture?

Group of people standing in front of building to take group picture during vacation at noon.

Here, we attempt to "narrate" the picture

CARE-05

Could you imagine picture of “people standing in front of building at noon”?

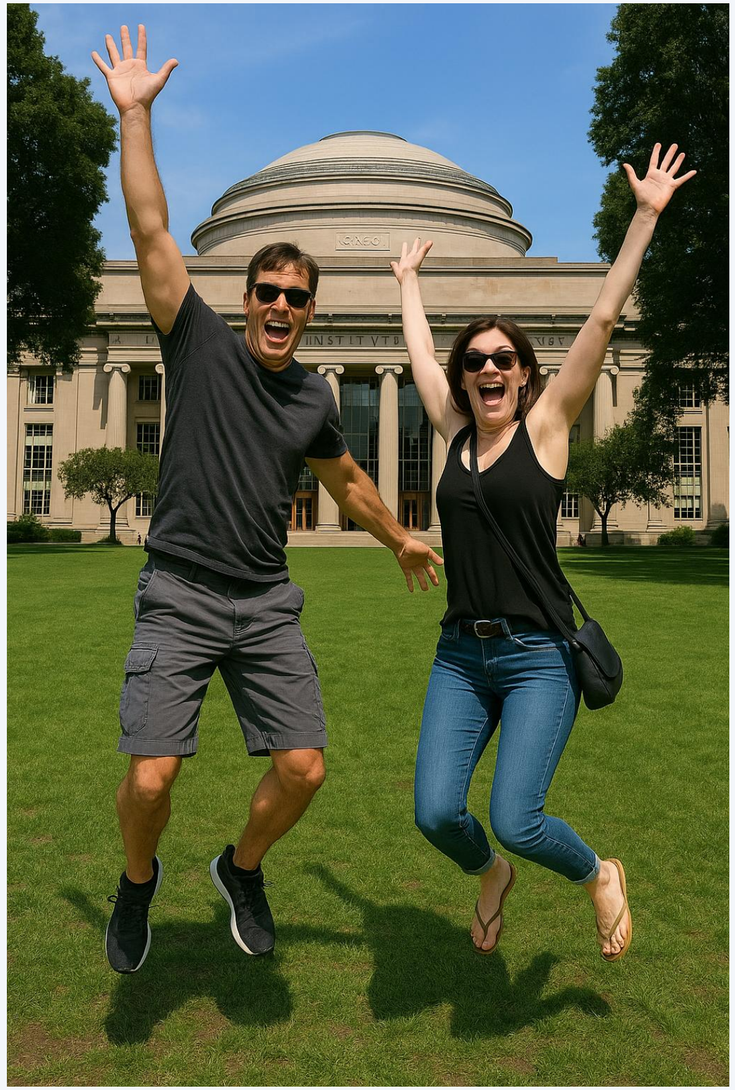
CARE-05

Here, we attempt to “generate” the picture

CARE-06

What will happen next?

CARE-06

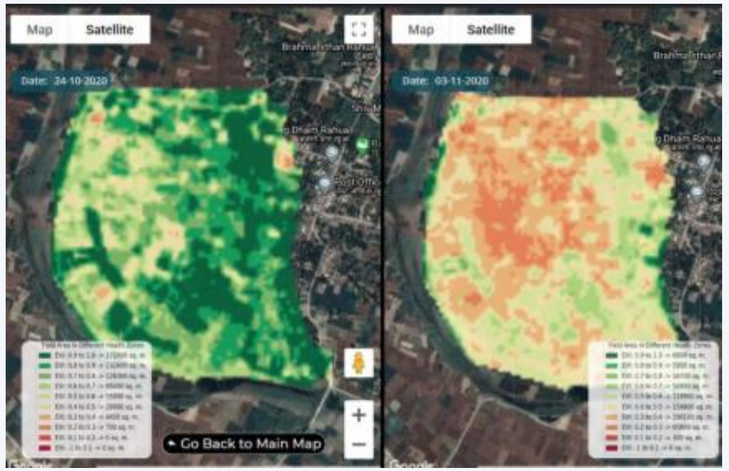


The children left and the parent take a group picture.

Here, we attempt to "predict/forecast" the next one.

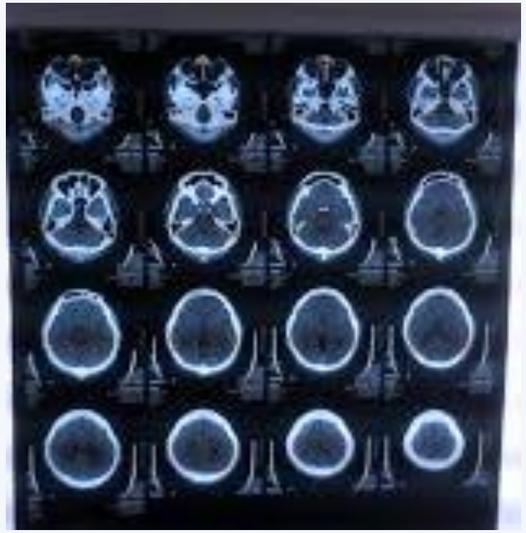
Real Implementation

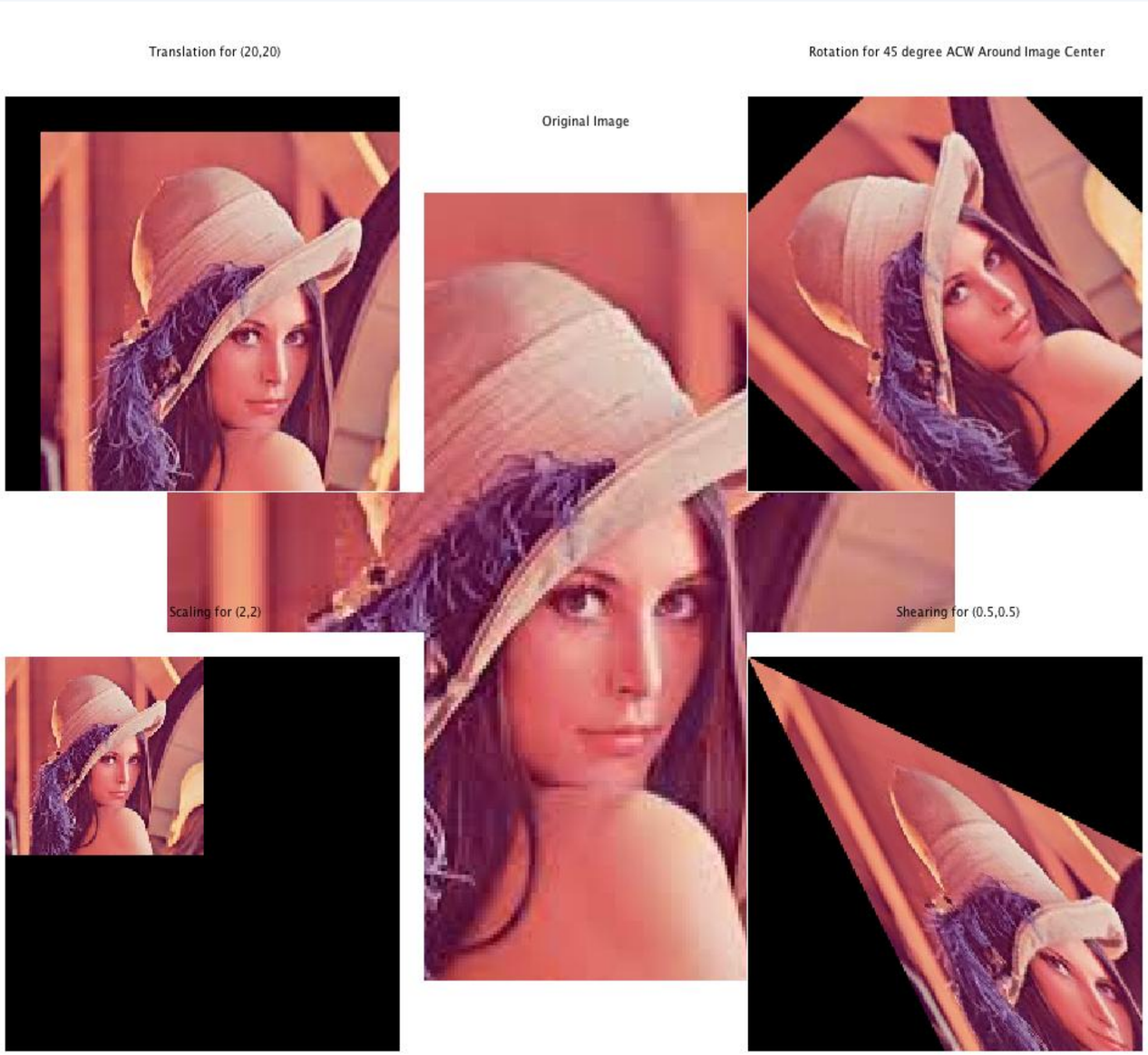
1. **Yield Estimation and Fruit Counting**
    1. Vision systems on robots or phones count apples, grapes, or wheat heads to estimate yield.



1.Tumor Detection in MRI and CT Scans

1. Medical systems use image segmentation to automatically **detect brain tumors**, lung nodules, or cancers in MRI/CT images.





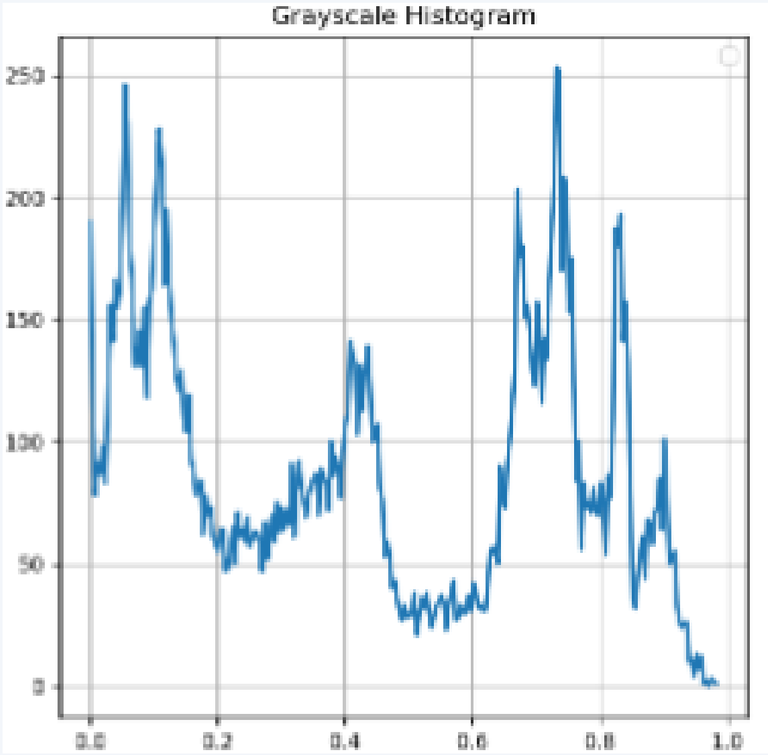
Tweaking Image...

scikit-image is a collection of algorithms for image processing. It is available free. (https://scikit-image.org/)

Play with Grayscale Image

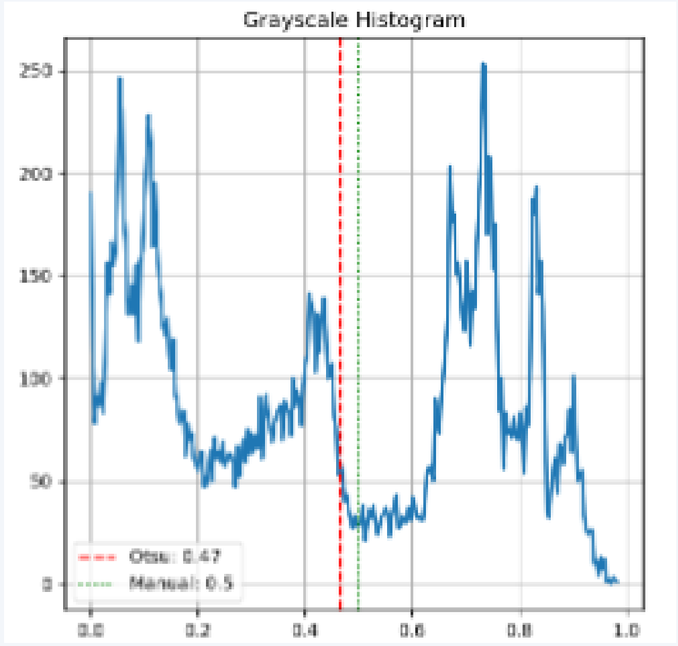
color.rgb2gray(*image*)

Histogram



What is the **histogram** say?
Any idea?

Remove the Background



Otsu Thresholding

Otsu's method finds the **optimal threshold value** that minimizes the **intra-class variance** (i.e., the variance within the foreground and background classes) or equivalently **maximizes the inter-class variance** (i.e., the variance between the two classes).

Manual Thresholding

Determine the threshold **manually**.

Resize, Shift, Flip, Rotate

Resize

Shifting

Flipping

Rotate

Convolution (Spatial Filtering)

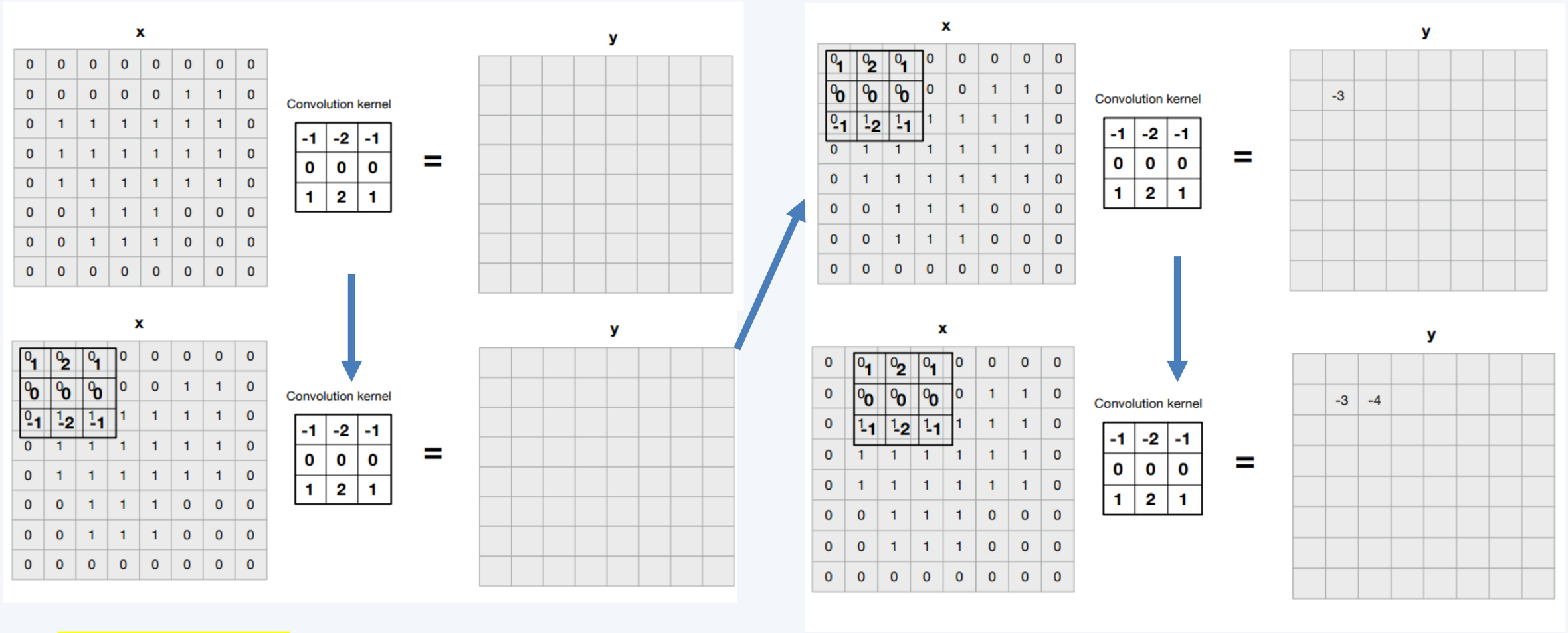
🎯 What is Convolution?

Convolution is a mathematical operation where a small matrix (called a **kernel** or **filter**) is passed over an image, modifying its pixel values based on their neighbors.

🧠 Intuition:

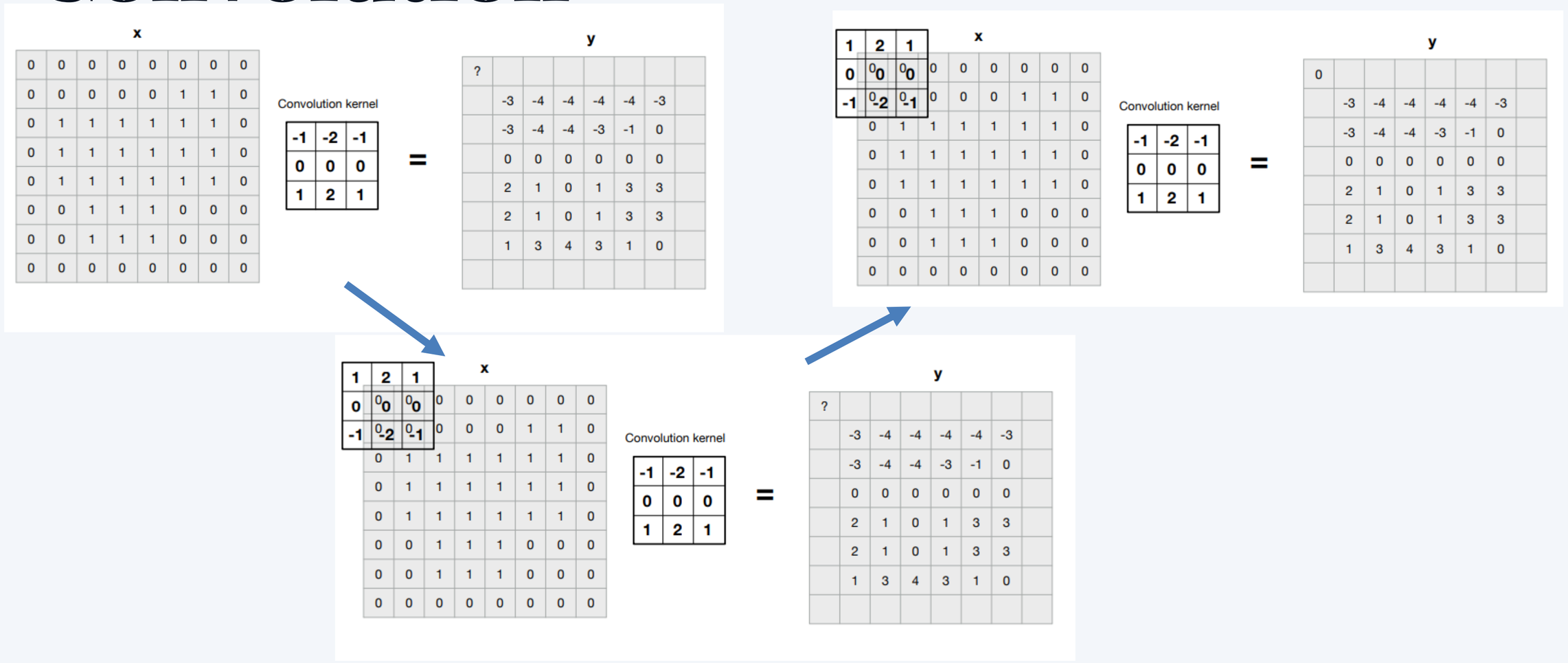
- Think of convolution like **looking at the image through a tiny window** (e.g. 3×3).
- For each pixel, we **compute a weighted sum** using the kernel and assign the result to the output pixel.

Convolution



Take attention! The kernel is inverted, and the result is sum of pair-wise multiplication.

Convolution



Take attention! The kernel is inverted, and the result is sum of pair-wise multiplication.

Convolution Example

Original

Blur

```
array([[0.11111111, 0.11111111, 0.11111111],
       [0.11111111, 0.11111111, 0.11111111],
       [0.11111111, 0.11111111, 0.11111111]])
```

Blurred

Sharpening

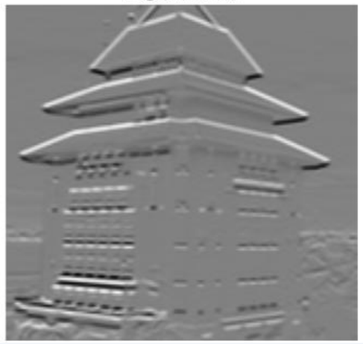
```
array([[ 0, -1,  0],
       [-1,  5, -1],
       [ 0, -1,  0]])
```

Sharpened

Edge Detection (Sobel)

```
array([[-1, -2, -1],
       [ 0,  0,  0],
       [ 1,  2,  1]])
```

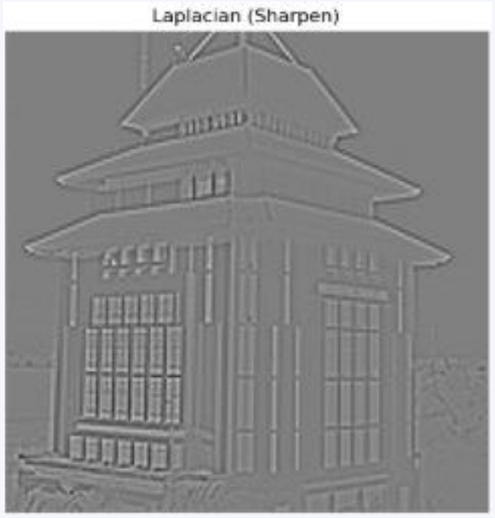
Edges (Sobel)



More Filtering...

Gaussian Filtering (Smoothing)

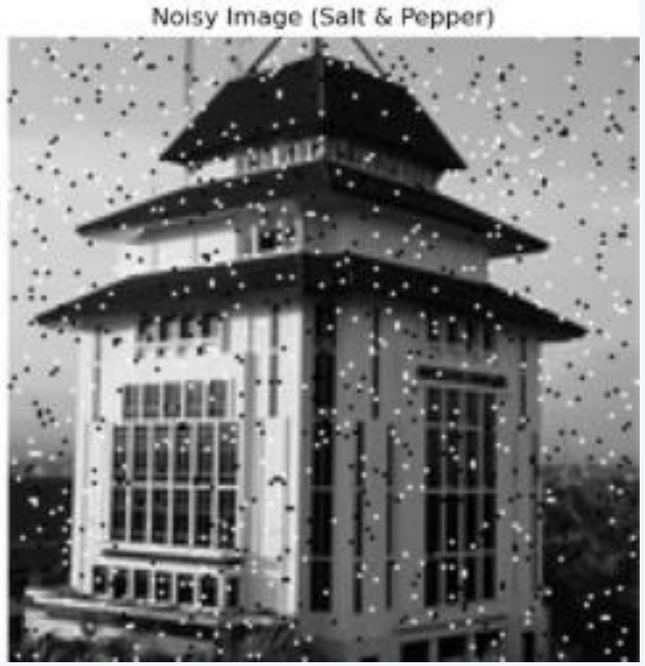
Laplacian



High-Pass Filter

More Filtering...

Noisy (Salt & Pepper)



Median Filtering

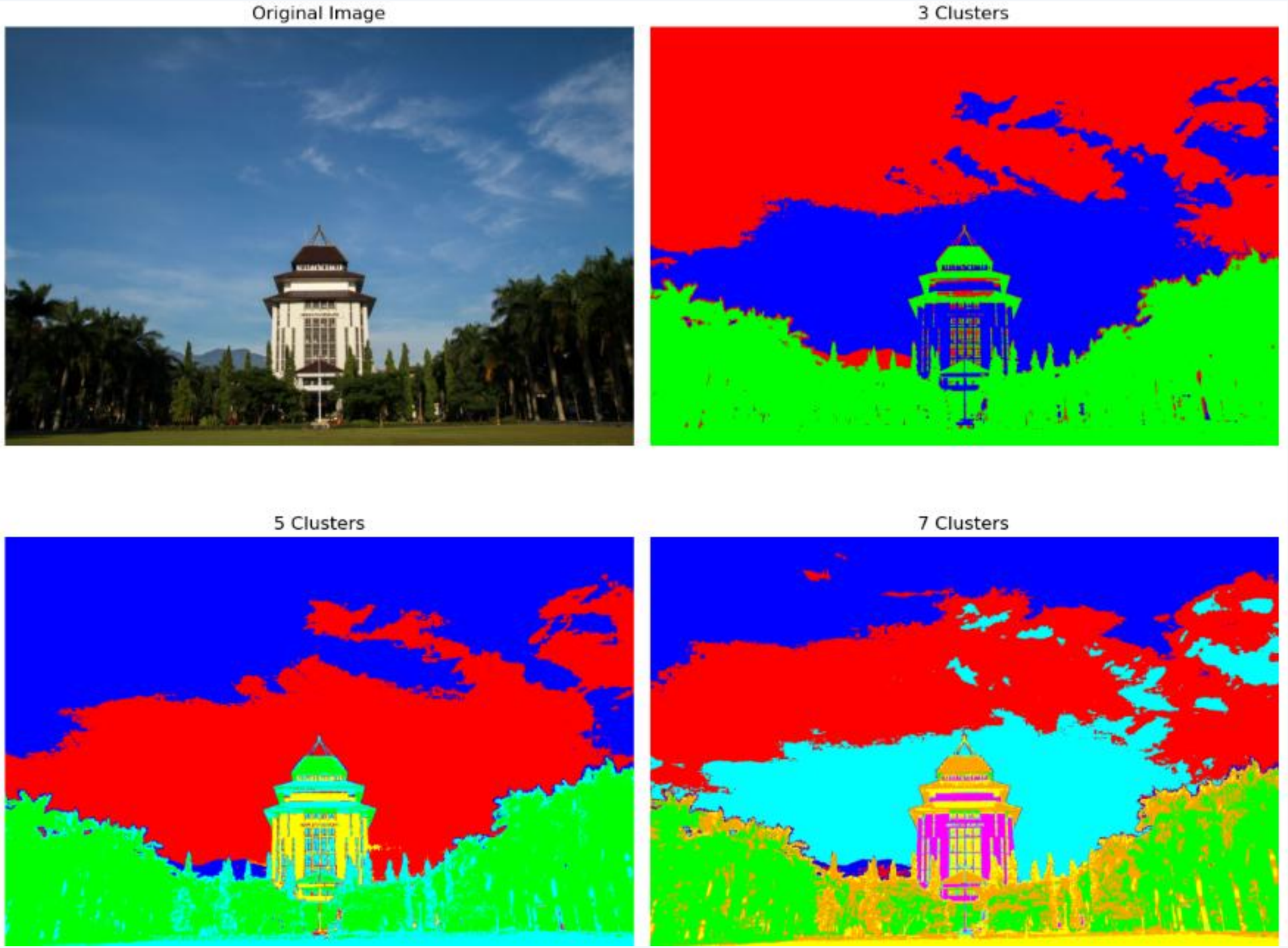
LET'S DO FUN

Let's try ***segmenting*** *the object using* ***simple clustering procedure (K-means)!***

PIPELINE

1. *Load the images*
2. *Create a data frame with each column represent the band*
3. *Do clustering using k-means (attempt using 3, 5, 7 cluster)*

RESULT



THANKS....

Real Application

150

150

Resolution: 150 x 150

Dilation and Erotion

Dilation
(Expand White Area)

Erotion
(Reduce White Area)

Compression

Compression can be made using the Fourier Transformation.
We change the data into Frequency domain and conduct the reconstructed one.