day - 1

Agenda :-

① Introduction to computer vision:-

② Image processing
                    ⤷ opencv, pillow

③ convolutional neural networks and different CNN architecture with practical hands on.

④ Transfer Learning and pre-trained Sot state of the art (SOTA) model.

⑤ object classification

⑥ object detection.

⑦ object segmentation.

⑧ keypoint detection / pose estimation.

⑨ Facial recognization

⑩ OCR → Optical character recognizer

⑪ YOLO, Detectron2, Tensorflow, Pytorch and so on.

⑫ Real world computer vision project implementation.

⑬ Deployment

(XIV) Latest Trends like vision Transformer, multimodel and so on.

(XV) object trancking

(XVI) GLANs

(XVII) Self driving ~~cat~~ car implementation

(XVIII) ~~~~ Defusion model

prerequisite

(I) Python programming → OOP

(II) Basic of DL → NN

→ What is computer vision?

It enables a machine or a computer to have a vision system close to humans.

Input → sensing device → interpreting device
(camera)   (processor)

↓

output

→ Three foundational steps of cv
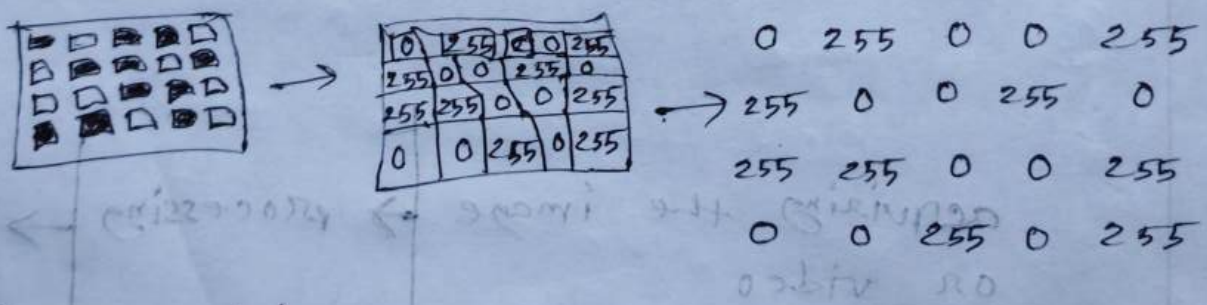
acquiring the image or video → processing → output

computer vision :

DL ⟶ NN

F ⟵ ① ANN ⟶ Artificial neural network
  ↳ tabular data

CV ↙ ② CNN ⟶ convolutional neural networks
  ↳ Image data, video data

NLP ↙ ③ Rnn ⟶ Recurrent neural network
  ↳ text / Audio data
       (time series data)

We work with CNN mainly.

so we need to,

Feature extraction:-

0	255	0	0	255
255	0	0	255	0
255	255	0	0	255
0	0	255	0	255

0 255 0 0 255
→255 0 0 255 0
255 255 0 0 255
0 0 255 0 255

⟶ Image data is a ~~pixel~~ m
matrix pixel value.
~~কোনো~~ object ~~এর~~ detection ~~করে~~ ~~করতে~~ হয়
~~তাই~~ feature লাগে,

→ CNN vs Human visual cortex

light →

thalamus

fig :- brain

visual cortex
↑ visual cortex
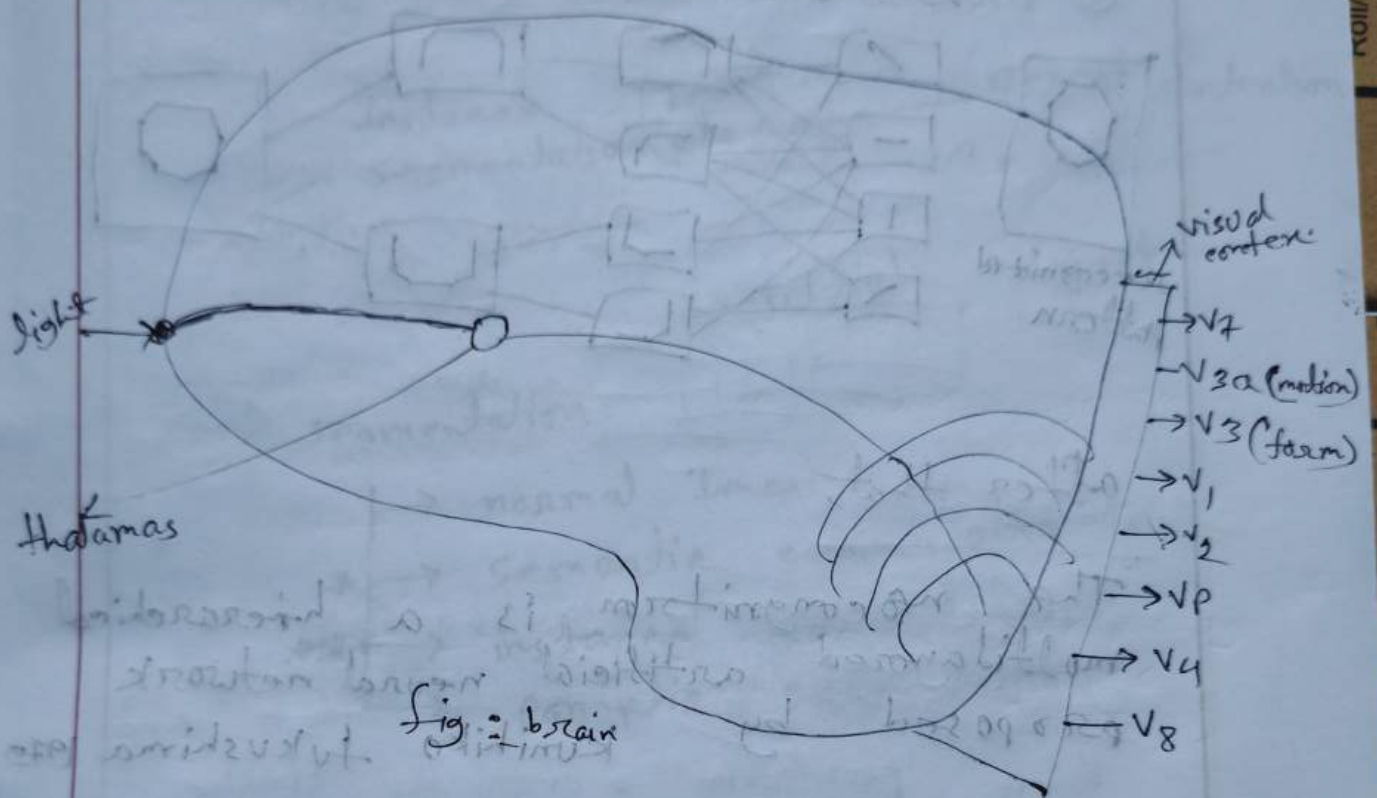→ V7
→ V3a (motion)
→ V3 (form)
→ V1
→ V2
→ VP
→ V4
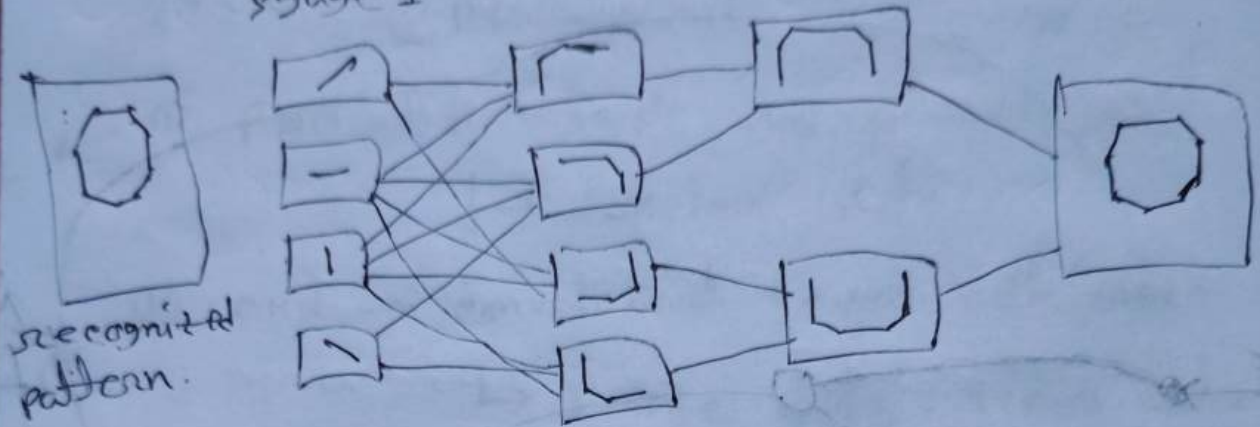→ V8

→ Hubel and wiesel :- Cat experiment

इलेक्ट्रोड
→ electrode

→ this research conclusion —

↳ simple cell ⟶ (blue) → orientation cell
↳ complex cell. ⟶ (Red) smaller receptive field
↳ bigere receptive field
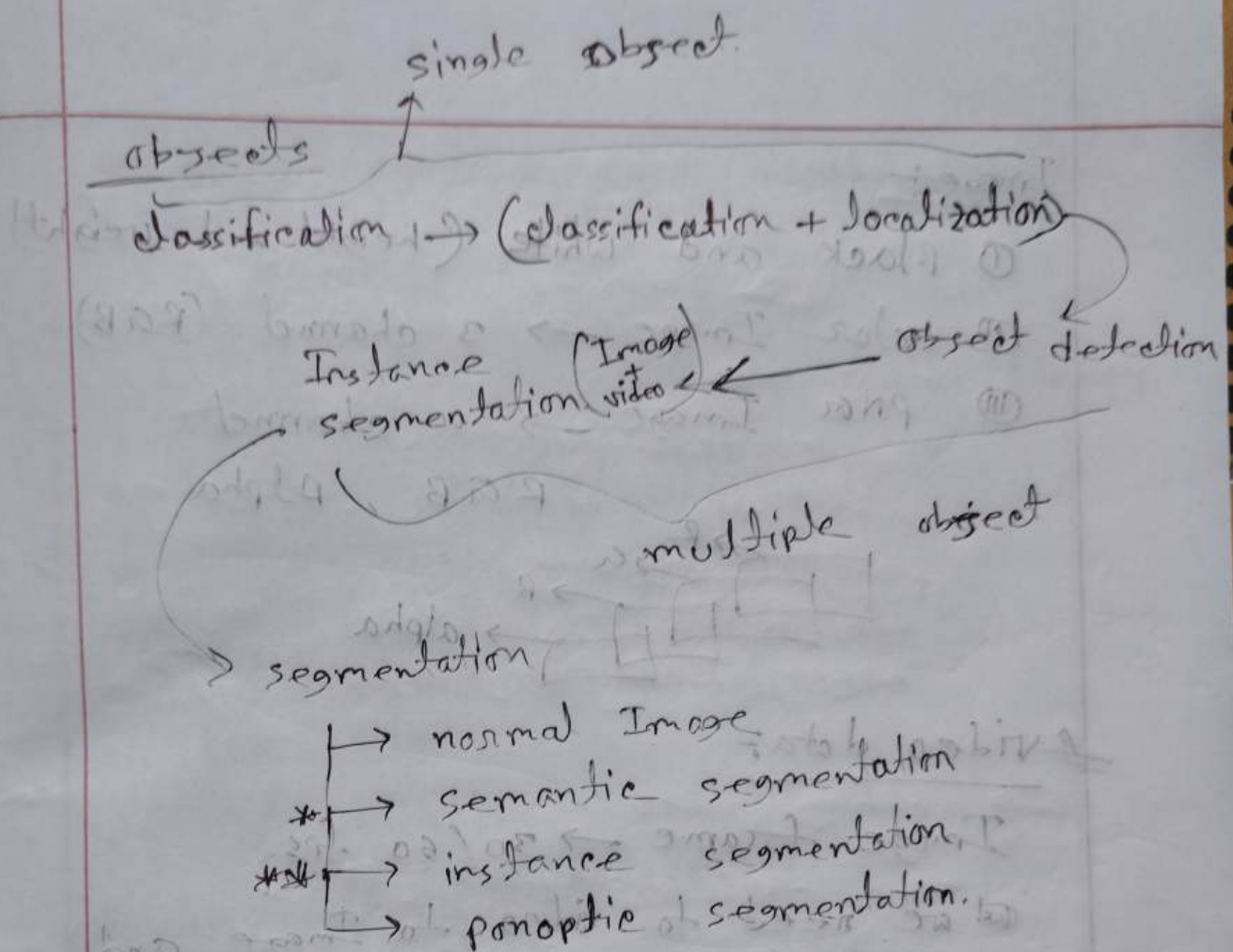
→ outcome :→

stage 1

recognital
pattern.

← after that,

The recognition is a hierarchical
multilayered artificial neural network
proposed by kunihiko fukushima 1979

after that:

"Gradient based learning applied
to Document recognition" → paper (yann
lecun

single object

objects
↑

classification → (classification + localization)

Instance
segmentation (Image + video) ← object detection

multiple object

→ segmentation

├→ normal Image
*├→ semantic segmentation
**├→ instance segmentation
└→ ponoptic segmentation.

application:-

 ① self driving car
 ② facial ~~nege~~ recognition
 ③ Detection system
 ④ AI → meta glass

challenges:-

 ① Data aquisation / qualitiful data
 ② hardware → GPU → A100 / V100
 ③ proper knowledge of domain.

Image :-

① Black and white (height and weight)

② color Image → 3 channel (RGB)

③ PNG Image → 4 channel

RGB , Alpha
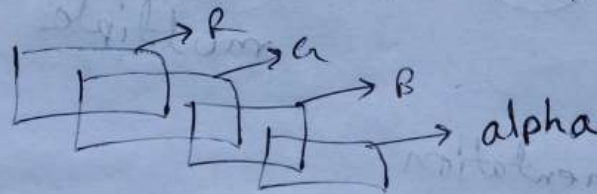
video data :

Image frame → 30/60 fps.

→ we need to video to Image and then perform over the image.