LOCATION EXTRACTION AND PEOPLE COUNTING IN VIDEOS

Guided By
Dr. Masilamani V

By
M Swetha
CED17I012

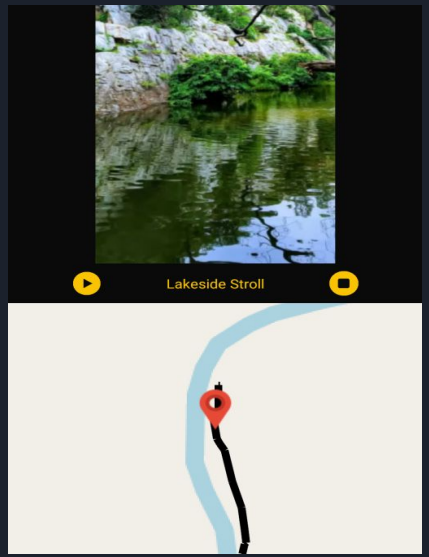
PROBLEM STATEMENT

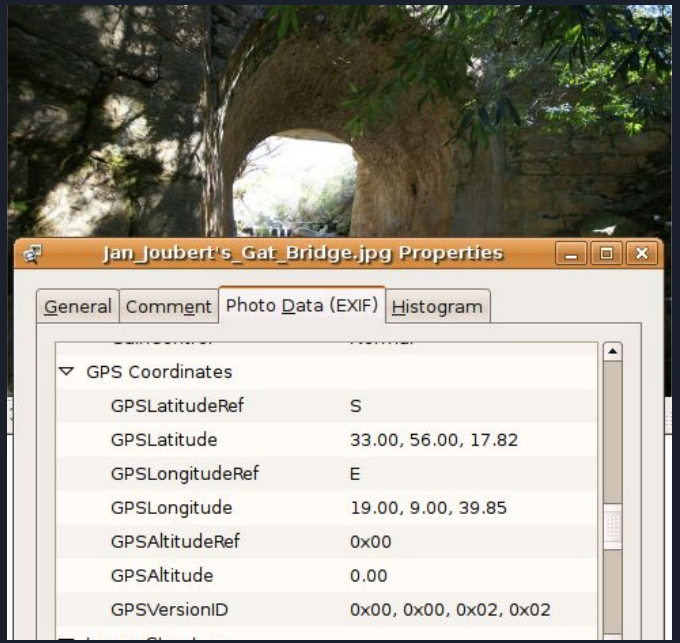
- To extract location in a video with location embedded in it and also find the number of people in the video frame by frame in a given GPS.

- To generate the application along with the implementation of embedding location on the video.

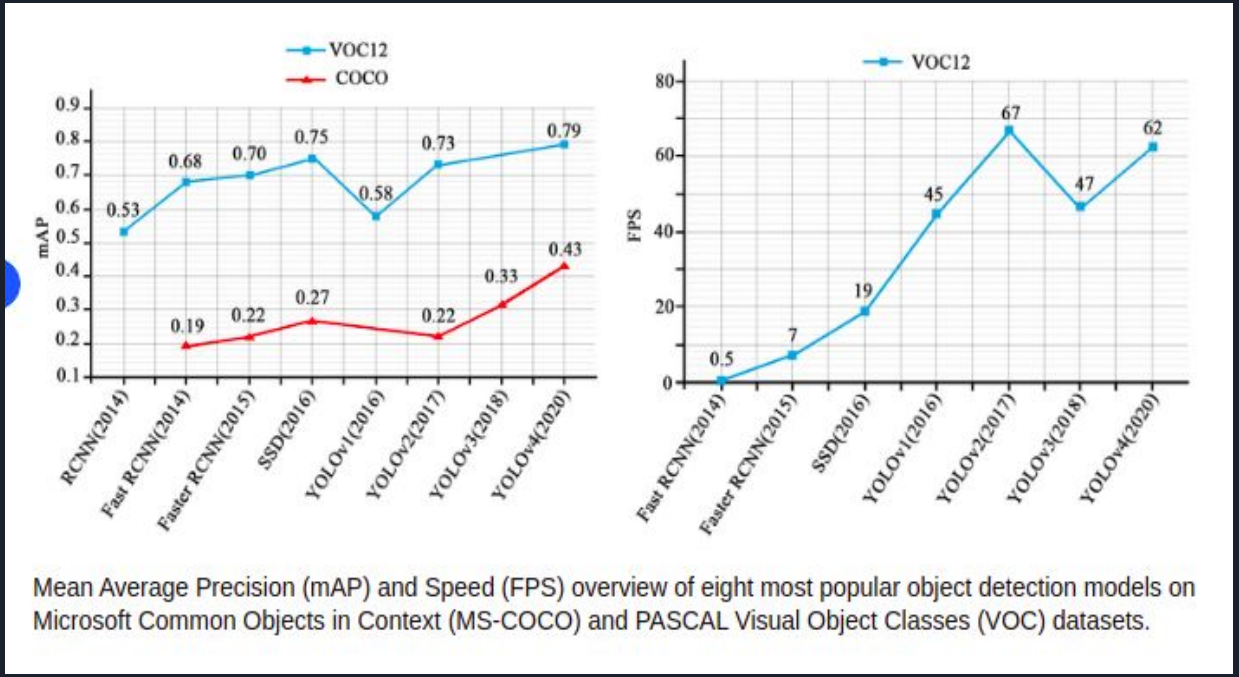
LITERATURE REVIEW

- For images, the geotagged data is embedded on the image as location data.
- Videos use displaying the google map while playing the video for location data.



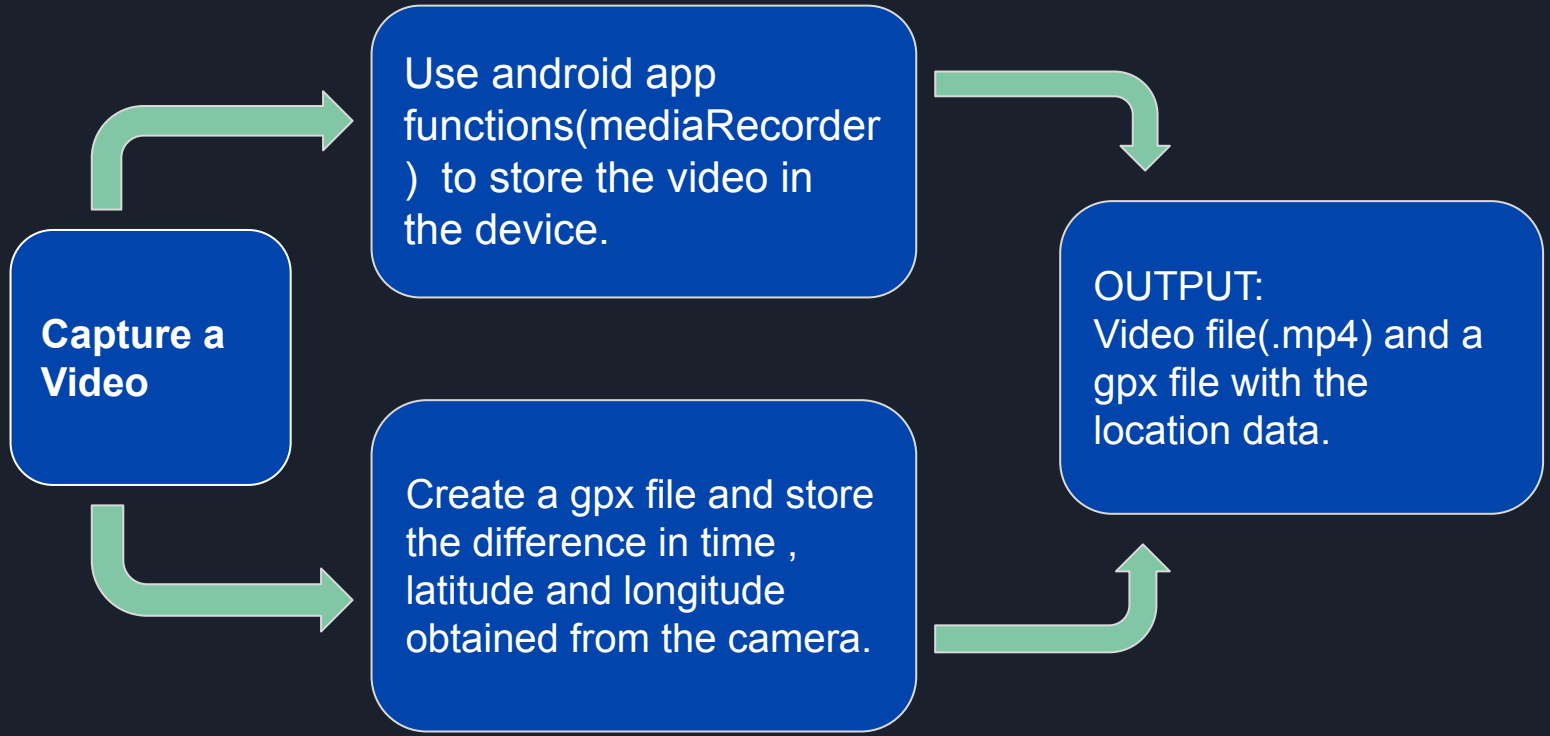


COMPARISON OF DETECTION ALGORITHMS

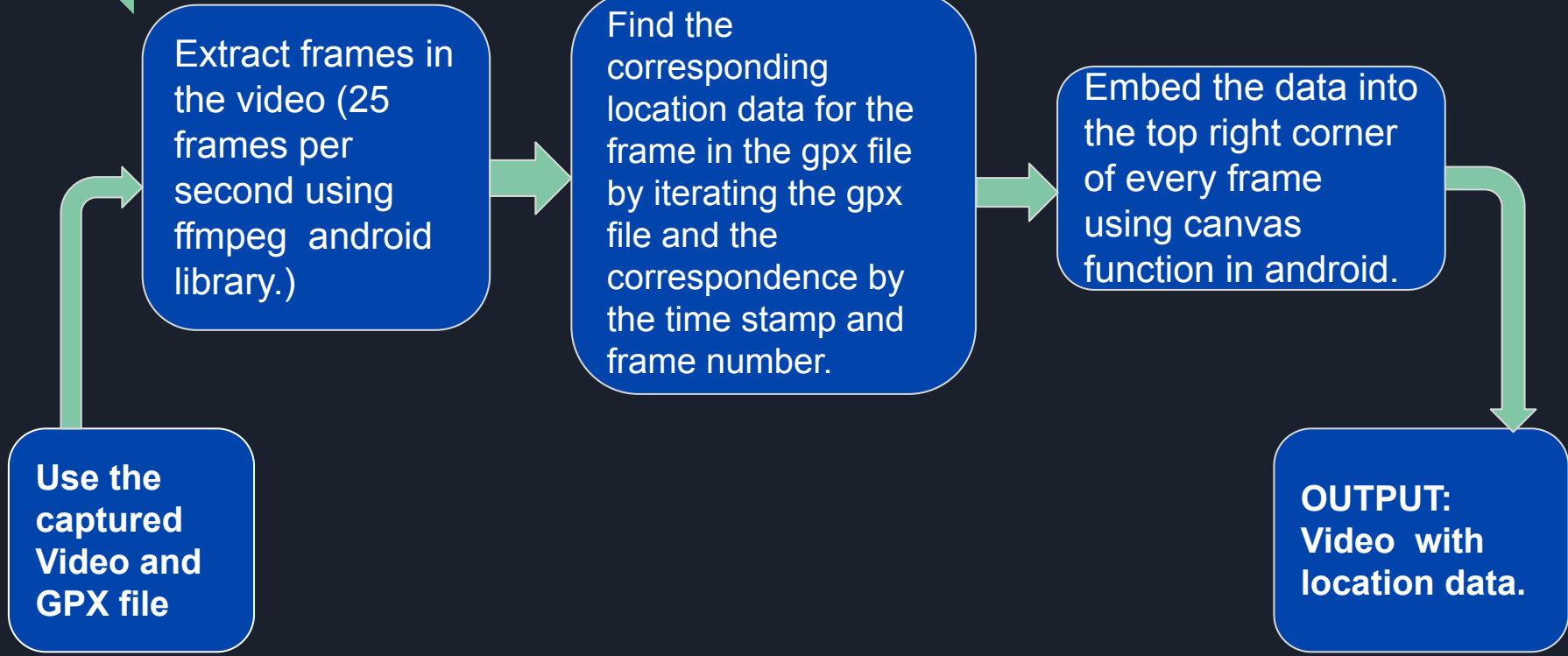


Mean Average Precision (mAP) and Speed (FPS) overview of eight most popular object detection models on Microsoft Common Objects in Context (MS-COCO) and PASCAL Visual Object Classes (VOC) datasets.

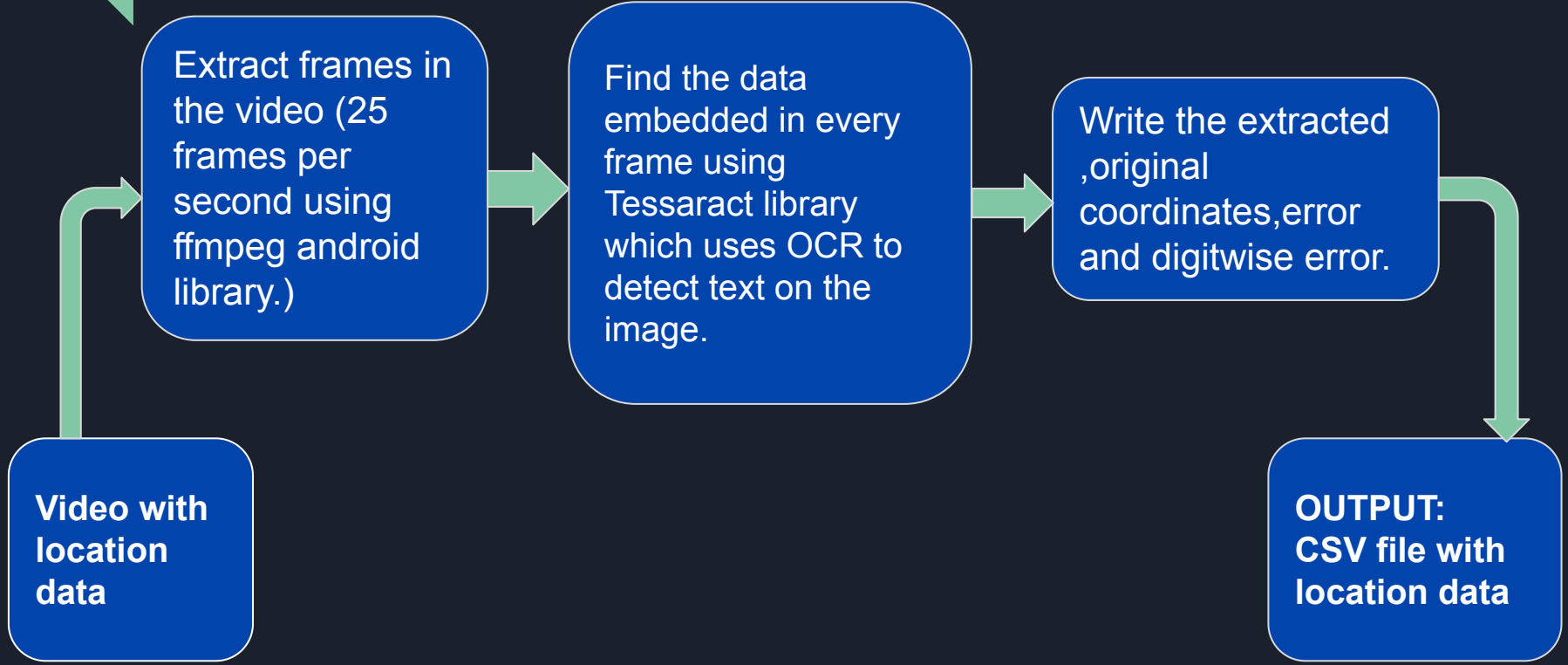
PROPOSED SOLUTION - LOCATION CAPTURE



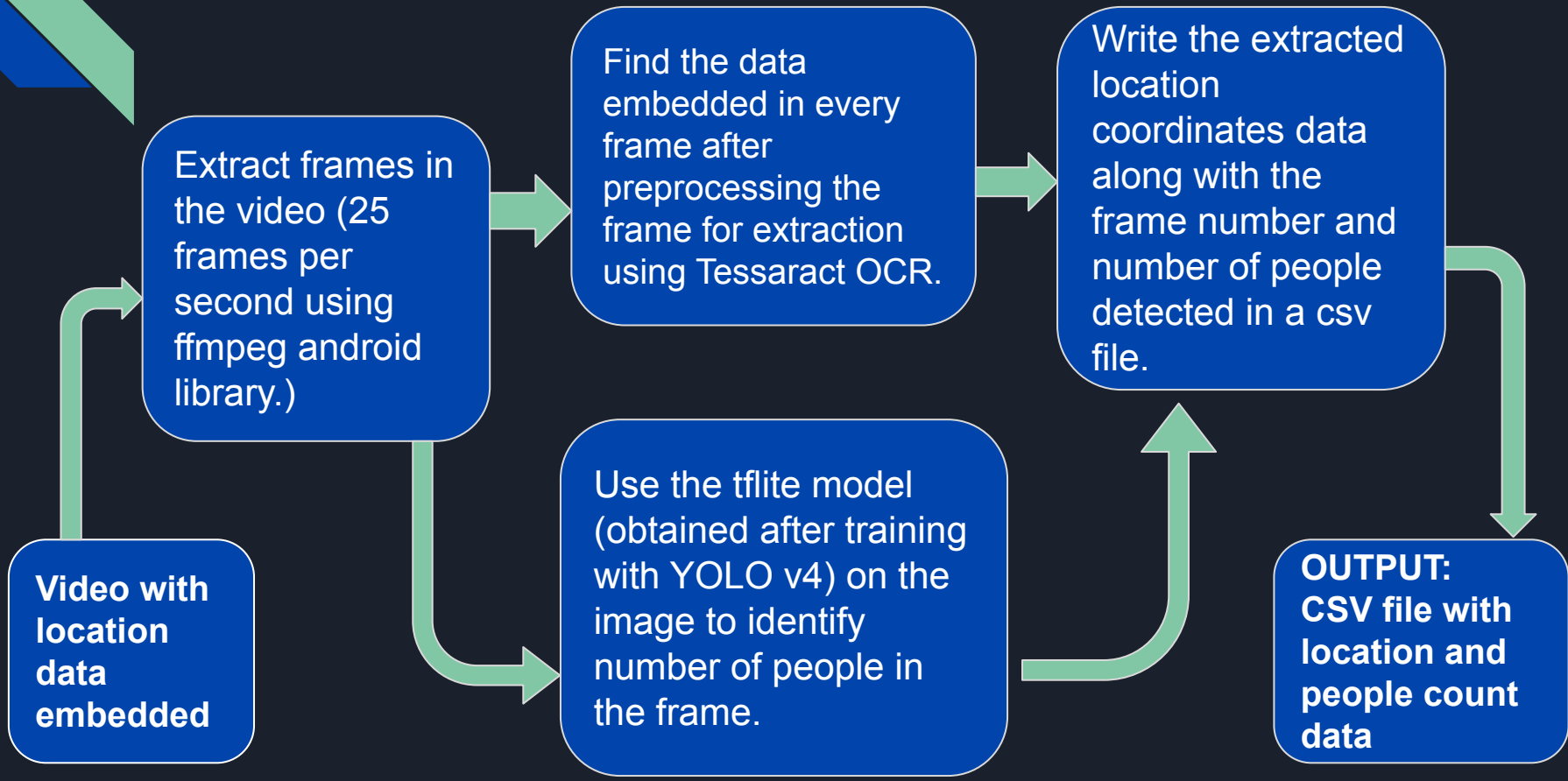
PROPOSED SOLUTION - EMBED LOCATION DATA



PROPOSED SOLUTION - EXTRACT LOCATION DATA AND FIND ACCURACY



PROPOSED SOLUTION - PEOPLE COUNTER

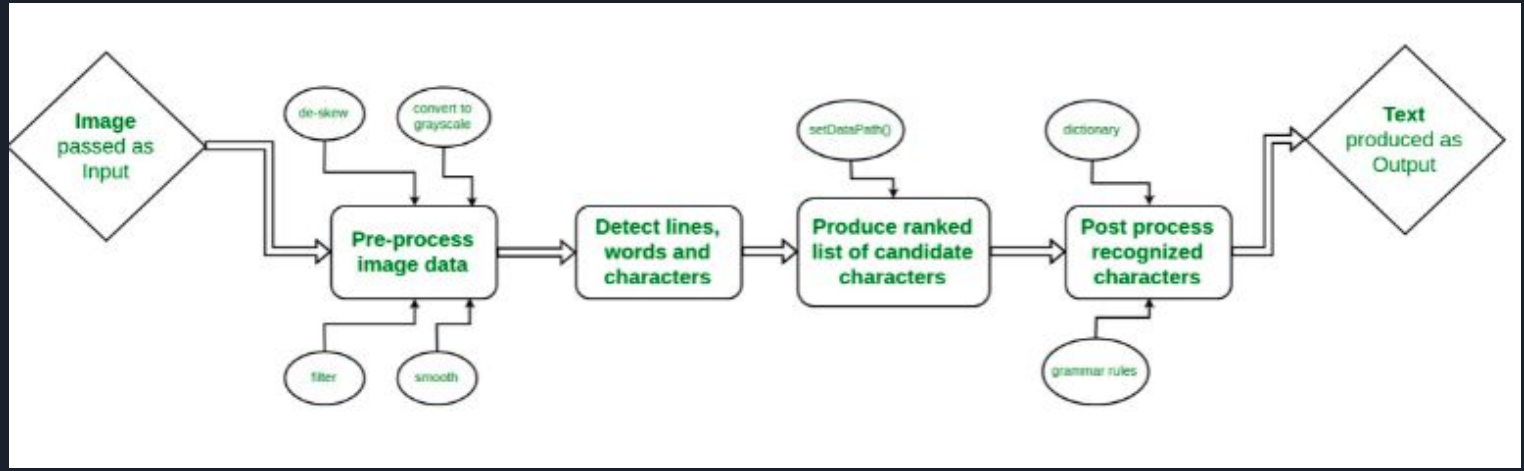


METHODOLOGY - STORE DATA ON THE FRAME

- Canvas is a class in android for 2D drawings which uses a 2D Graphics Library.
- To write on the image, the bitmap is first drawn on to a canvas.
- We can overwrite the drawn pixels with the pixels of the desired shape.
- This is used for making the background black for a certain width and writing the text as a drawing on the image.

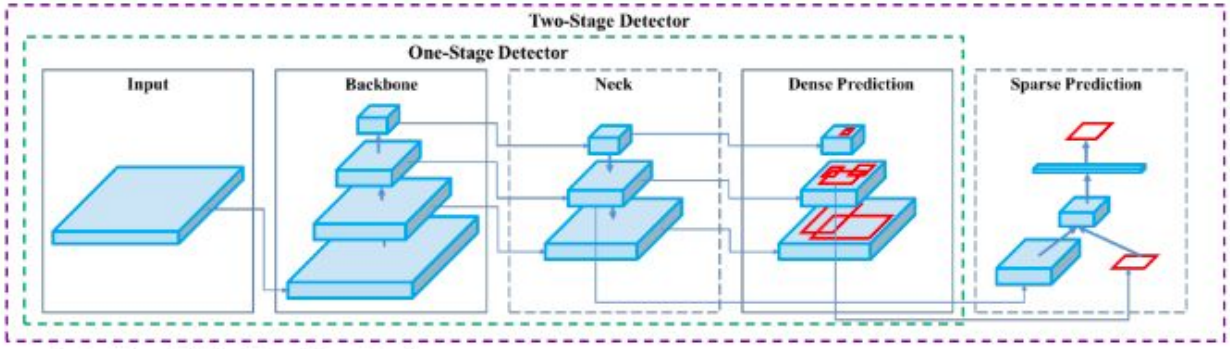
METHODOLOGY - RETRIEVE DATA FROM THE FRAME

Optical Character Recognition is the technology used to retrieve text from image.



METHODOLOGY - COUNTING PEOPLE IN THE VIDEO

YOLO v4 is used for counting people in video frame by frame.



Input: { Image, Patches, Image Pyramid, … }

Backbone: { VGG16 [68], ResNet-50 [26], ResNeXt-101 [86], Darknet53 [63], … }

Neck: { FPN [44], PANet [49], Bi-FPN [77], … }

Head:

Dense Prediction: { RPN [64], YOLO [61, 62, 63], SSD [50], RetinaNet [45], FCOS [78], … }

Sparse Prediction: { Faster R-CNN [64], R-FCN [9], … }

PRE-PROCESSING FOR IMPROVING TEXT DETECTION

- Image Cropping : Reduced Image Size, Removal of unwanted Characters
- Recognize only digits: Reduced misclassifications.
- Image Segmentation

RESULTS - OCR Algorithm

Sample Number	Accuracy Before processing	Accuracy After processing
1	100	100
2	92.89	93.1
3	89.66	94.61
4	86.8	95.20
5	93.8	92.9

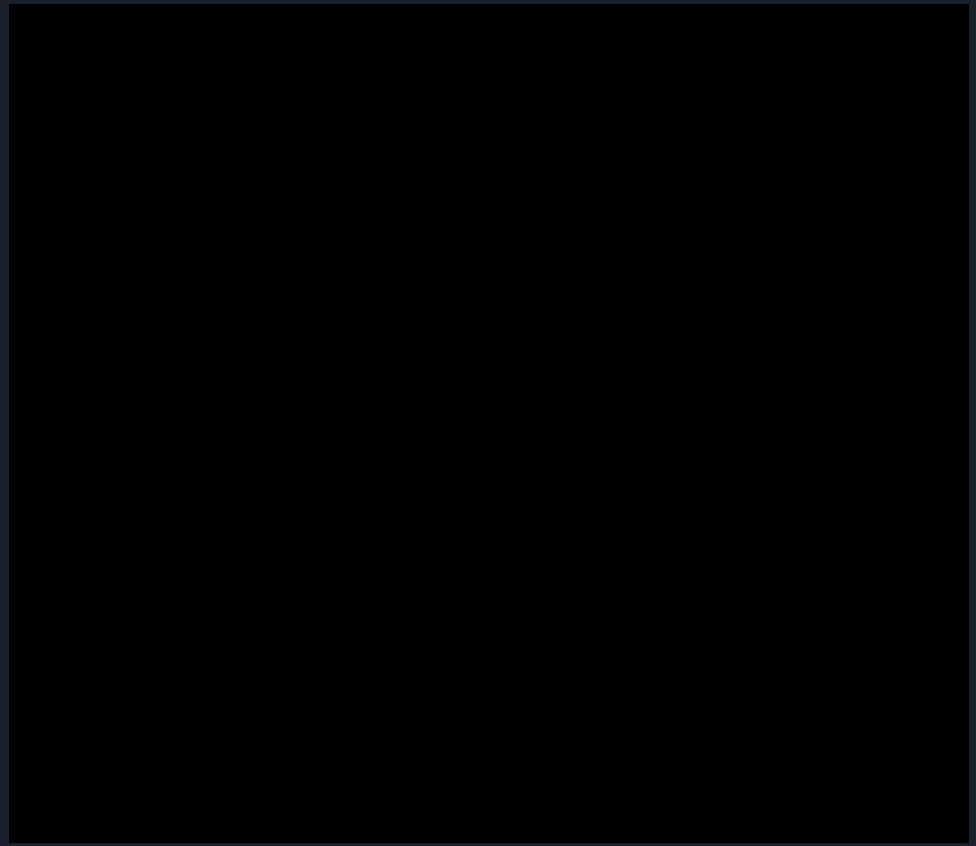
Average Accuracy Before Processing : **92.63**
Average Accuracy After Processing : **95.2**

Digit wise accuracy After Processing

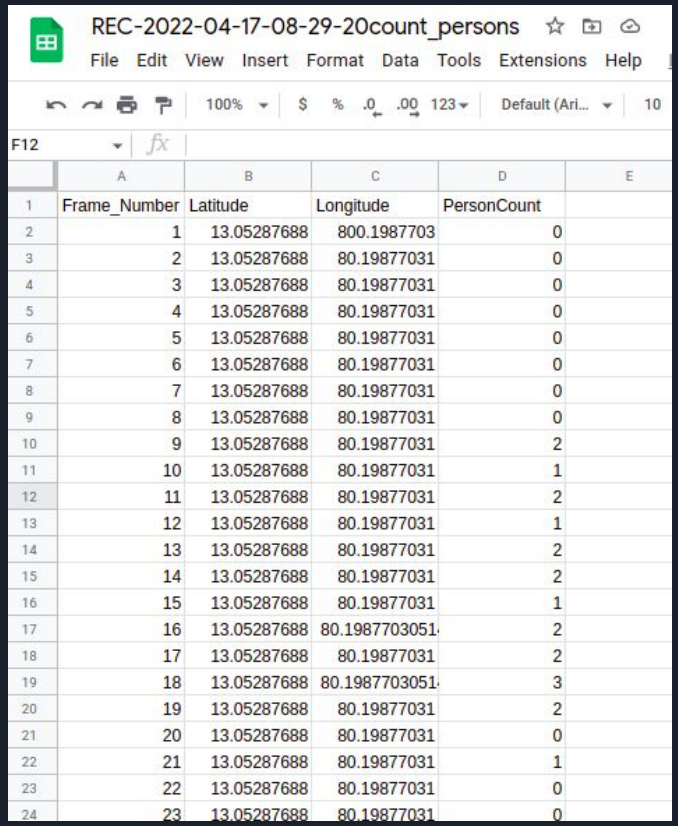
Digits	Number of Misclassifications	Total Number of Digits	Accuracy
0	140	8837	98.41575195
1	216	8518	97.46419347
2	259	4832	94.63990066
3	247	7705	96.79428942
4	217	8355	97.40275284
5	287	4238	93.22793771
6	181	5593	96.76381191
7	173	3823	95.47475804
8	236	7598	96.89391945
9	190	8944	97.87567084

RESULTS - DEMO TO EMBED VIDEOS

RESULTS - DEMO FOR PEOPLE COUNTER



RESULTS - CSV FILE GENERATED



REC-2022-04-17-08-29-20count_persons

File Edit View Insert Format Data Tools Extensions Help

100% $ % .0 .00 123 Default (Ari... 10

F12

	A	B	C	D	E
1	Frame_Number	Latitude	Longitude	PersonCount	
2	1	13.05287688	800.1987703	0	
3	2	13.05287688	80.19877031	0	
4	3	13.05287688	80.19877031	0	
5	4	13.05287688	80.19877031	0	
6	5	13.05287688	80.19877031	0	
7	6	13.05287688	80.19877031	0	
8	7	13.05287688	80.19877031	0	
9	8	13.05287688	80.19877031	0	
10	9	13.05287688	80.19877031	2	
11	10	13.05287688	80.19877031	1	
12	11	13.05287688	80.19877031	2	
13	12	13.05287688	80.19877031	1	
14	13	13.05287688	80.19877031	2	
15	14	13.05287688	80.19877031	2	
16	15	13.05287688	80.19877031	1	
17	16	13.05287688	80.1987703051	2	
18	17	13.05287688	80.19877031	2	
19	18	13.05287688	80.1987703051	3	
20	19	13.05287688	80.19877031	2	
21	20	13.05287688	80.19877031	0	
22	21	13.05287688	80.19877031	1	
23	22	13.05287688	80.19877031	0	
24	23	13.05287688	80.19877031	0	

CHALLENGES

- Ensuring that the location data is captured correctly for every internal frame.
- Optimising the code to ensure minimum image processing time apart from the detection and extraction algorithms.

NOVELTY

- Writing GPS on the image such that OCR works efficiently
- Incorporating some pre-processing techniques to ensure better accuracy in extracting the OCR from the video.

USE CASE

- To find the number of people on the given location - Traffic Congestion.

- Video or image correspondence of the patient by embedding the patient details in case of medical data.

- Crowd Analysis in a political rally or movement of people in the railway station.

CONCLUSION AND FUTURE SCOPE

- The given video was successfully embedded with location data and detection of people along with Location extraction was completed with the optimal accuracy.

- Improving the performance of the application as a whole by further improving the speed and accuracy of the algorithms used.

REFERENCES

- Simple Location Logging Application
- YOLO v4 Algorithm Research Paper

GITHUB LINK AND SAMPLES

- Github Code of the application
- Drive Link to video Samples tried and tested to improve the algorithm