

MidPrep Problem Statement

CloudPhysician: The Vital Extraction Challenge

INTRODUCTION

- **Objective** - To extract the data for the vitals like Heart Rate, SPO2, RR, SBP, DBP, MAP from the monitor screen of patients
- **Background** - This problem statement is a part of CloudPhysician's pursuit of creating a Smart ICU Solution
- Entire problem statement divided into 3 main tasks - monitor screen extraction, feature extraction and text recognition
- Graph digitization - to extract information from waveform based vitals

DATASET

- Provided with **3 types of datasets**:
 - Monitor Dataset
 - Classification Dataset
 - Unlabelled Dataset
- **Monitor dataset -**
 - 2000 images and coordinates of the monitor screen
- **Classification dataset -**
 - 4 folders of 250 images each and 4 data files with coordinates for features
 - Deal with null values
- **Unlabelled dataset -**
 - 7000 images of monitors without any coordinates for screen or features



	A	B
1	image_name	points
2	medicakolkata_mau	[459.5, 169.5, 886.0
3	kamalnayanbajaj_m	[422.4, 158, 967.8, 1
4	kamalnayanbajaj_m	[497.4, 184.9, 792.7
5	medicakolkata_ccu2	[353.799987792968
6	medicakolkata_ccu2	[246.9, 155.3, 756.5
7	hcgbhavnagar_micu	[228.9, 110.4, 808.4
8	medicakolkata_ccu1	[467.4, 188.1, 854.1
9	medicakolkata_ccu2	[337, 165.80000305
10	medicakolkata_ccu2	[238.4, 76.5, 721.59
11	medicakolkata_ccu1	[339.600006103515
12	medicasiliguri_icu_r	[398.4, 309.1, 665.8
13	medicakolkata_ccu2	[319.6, 122.8, 741, 1
14	medicakolkata_ccu3	[295.5, 137.3, 822.5
15	medicakolkata_ccu3	[443.200012207031

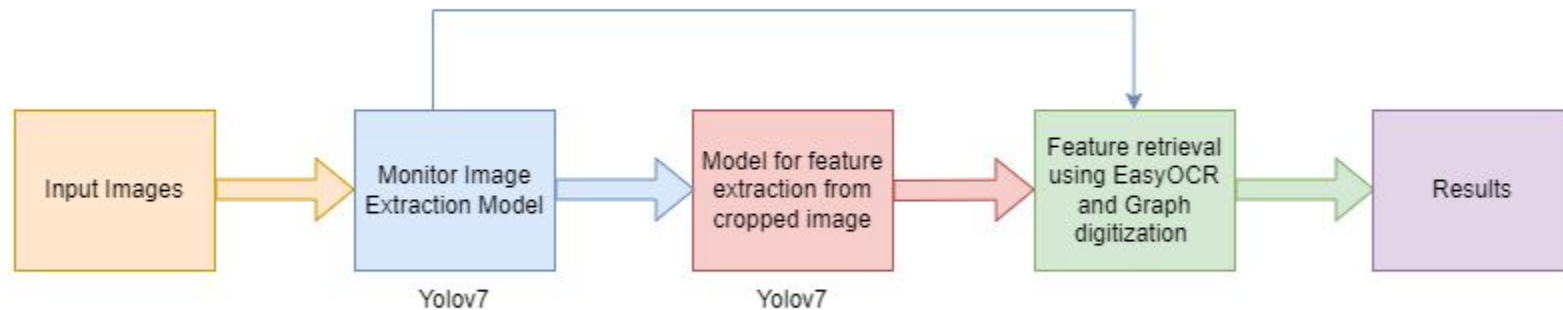
MONITOR DATASET

NULL VALUES			
SET A	0	0	0
	0	0	0
	1	1	1
	2	2	2
	4	4	4
	4	4	4
SET B	0	0	0
	1	1	1
	137	137	137
	137	137	137
	137	137	137
	137	137	137
SET C	0	0	0
	6	6	6
	34	34	34
	41	41	41
	14	14	14
	15	15	15
SET D	0	0	0
	1	1	1
	2	2	2
	14	14	14
	71	71	71
	71	71	71
ENTIRE SET			



CLASSIFICATION DATASET

MAIN PIPELINE



MODEL M1 - Monitor Screen Extraction

- **Training** - 1800 images, **Validation** - 200 images, **35 epochs**
- Uses **Yolov7** as the object detection framework
- **Input** - Monitor images from dataset with a label file containing coordinates for the screen
- **Output** - Tensor containing bounding box coordinates and confidence score
- Accurately detects the screen from the given image forming the basis for the next stage of the pipeline

MODEL M2 - Feature Extraction

- **Training** - 800 images, **Validation** - 200 images, **45 epochs**
- Uses **Yolov7** for object detection
- **Input** for Yolov7 model - image and label file containing coordinates for bounding boxes
- **Output** - Tensor containing bounding box coordinates and confidence score for each feature on the screen
- Further, through text recognition, the features are extracted
- Graph digitization is applied to extract information from the graphs on the monitor

Text Recognition

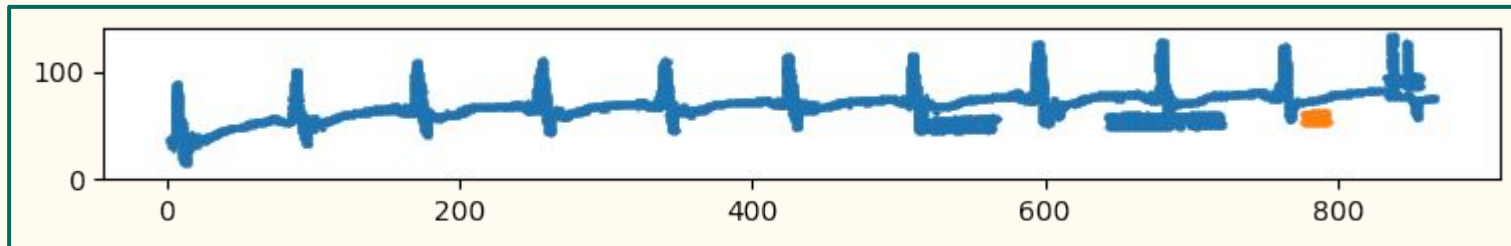
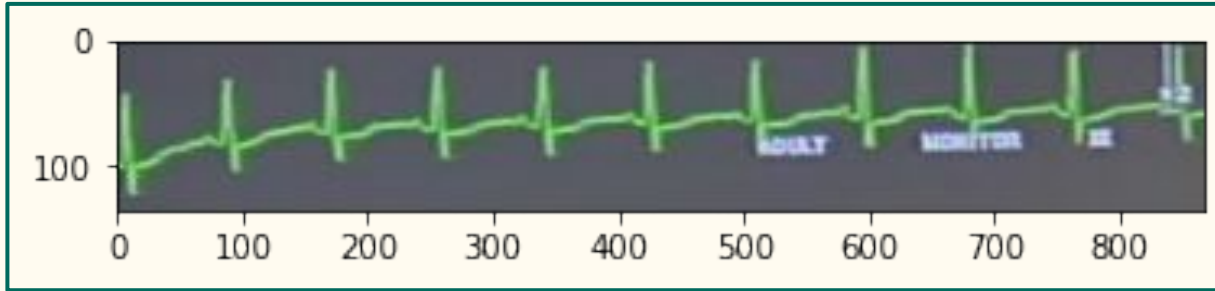
- Based on the dataframe containing bounding box coordinates, feature images are cropped and text is extracted using **EasyOCR**
- Remove non-numeric characters from the predicted values for vitals using clean string



Example of cropped image containing a bracket

Graph Digitization

- Get insights from the graph features on the monitor
- Use of **k-means clustering**, **thresholding**, **masking**, **DBSCAN clustering**



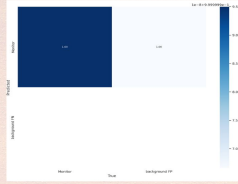
RESULTS

➤ Inference Time :

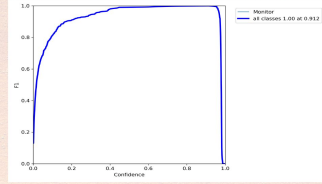
- Ran on Google Colab Notebooks
- Total inference time is around **3.5 s**
- This includes preprocessing for images, feature extraction and prediction

➤ Model Wise Results :

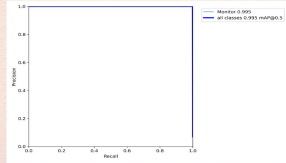
- Model M1
 - F1 score approaching **1** at a confidence of **0.912**
- Model M2
 - F1 score of about **0.93** at a confidence of **0.702**



Confusion Matrix

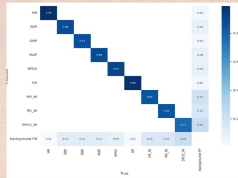


F1 Curve

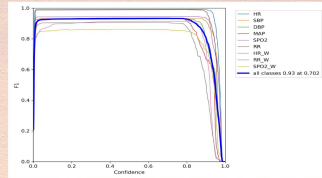


PR Curve

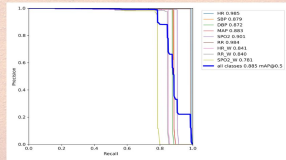
MODEL M1 - MONITOR SCREEN EXTRACTION



Confusion Matrix



F1 Curve



PR Curve

MODEL M2 - FEATURE EXTRACTION

THANK YOU