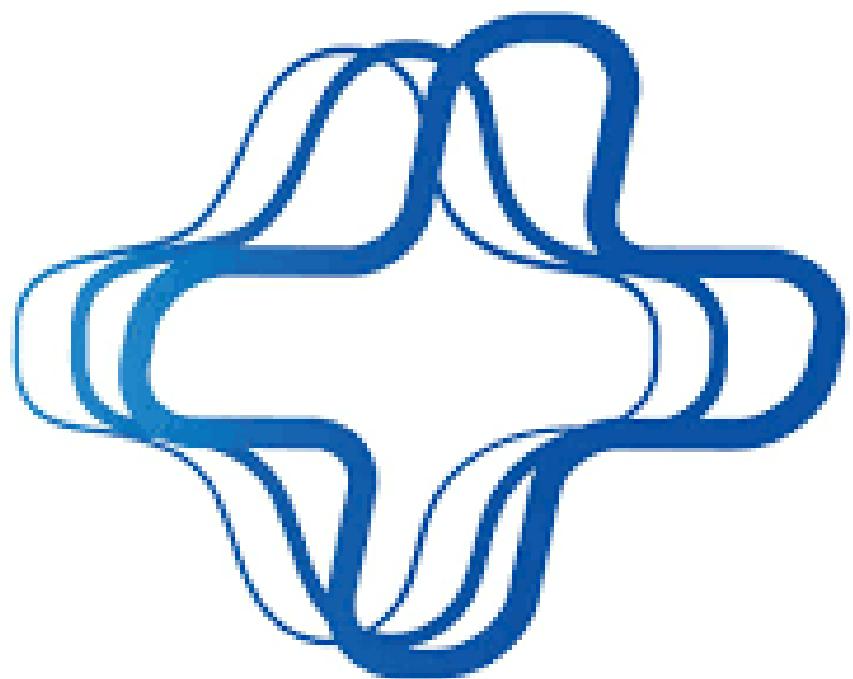


Team 43

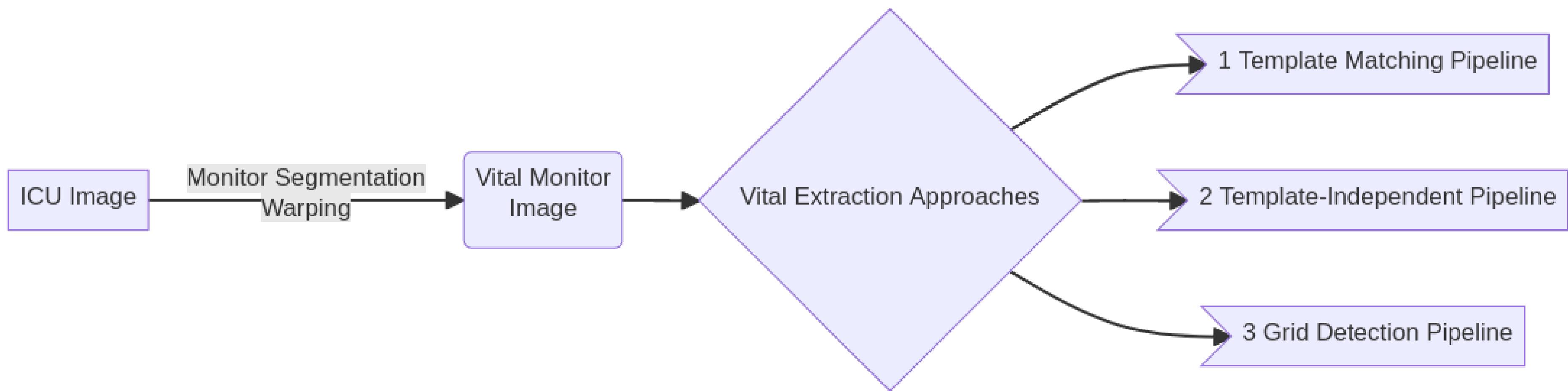


INTER IIT TECH MEET 11.0

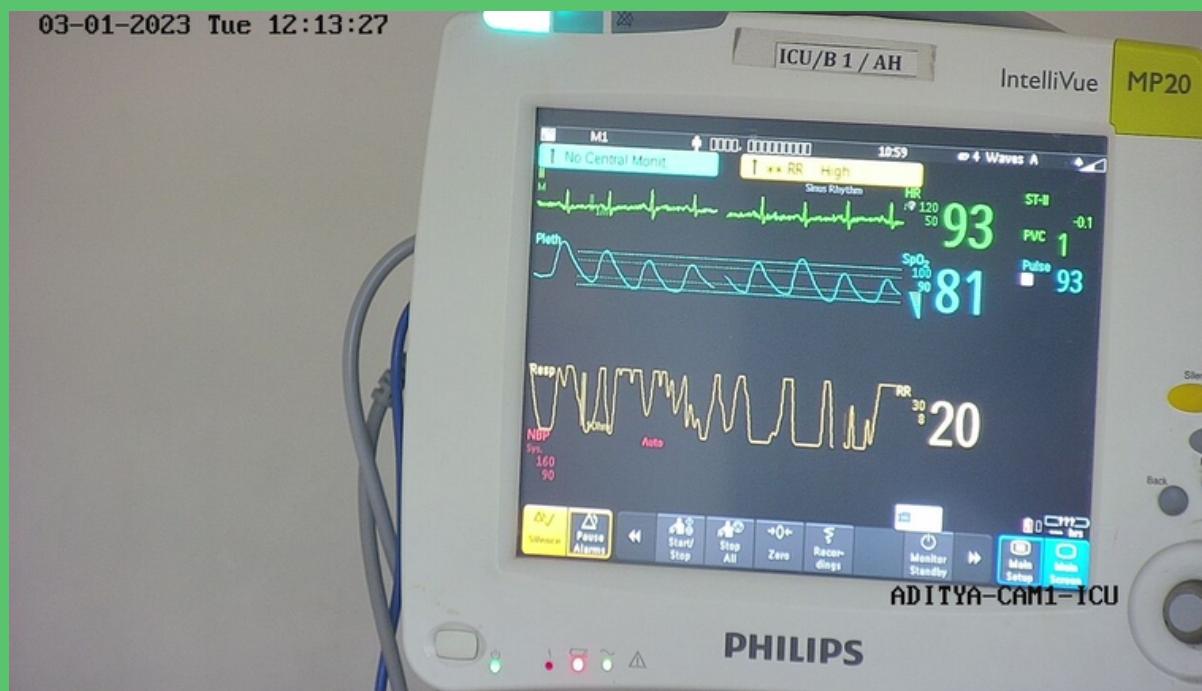
THE VITAL EXTRACTION CHALLENGE

Cloudphysician

ROADMAP



MONITOR SEGMENTATION

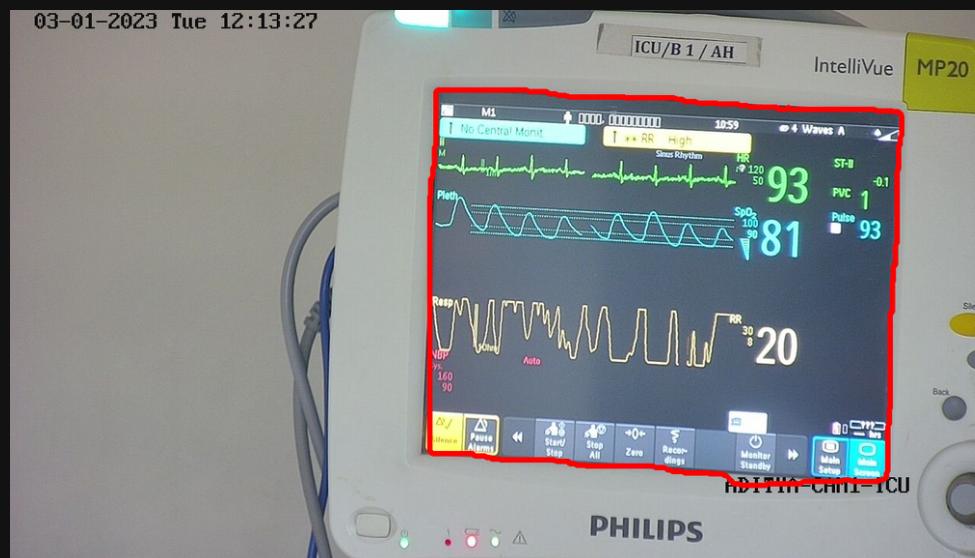


MONITOR SEGMENTATION

01

Screen Segmentation

Identify the boundaries of the screen from the surroundings



02

Corner Identification

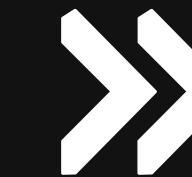
Identify the corners of the segmented screen, convert polygonal bounding box to quadrilateral bounding box



03

Image Warping

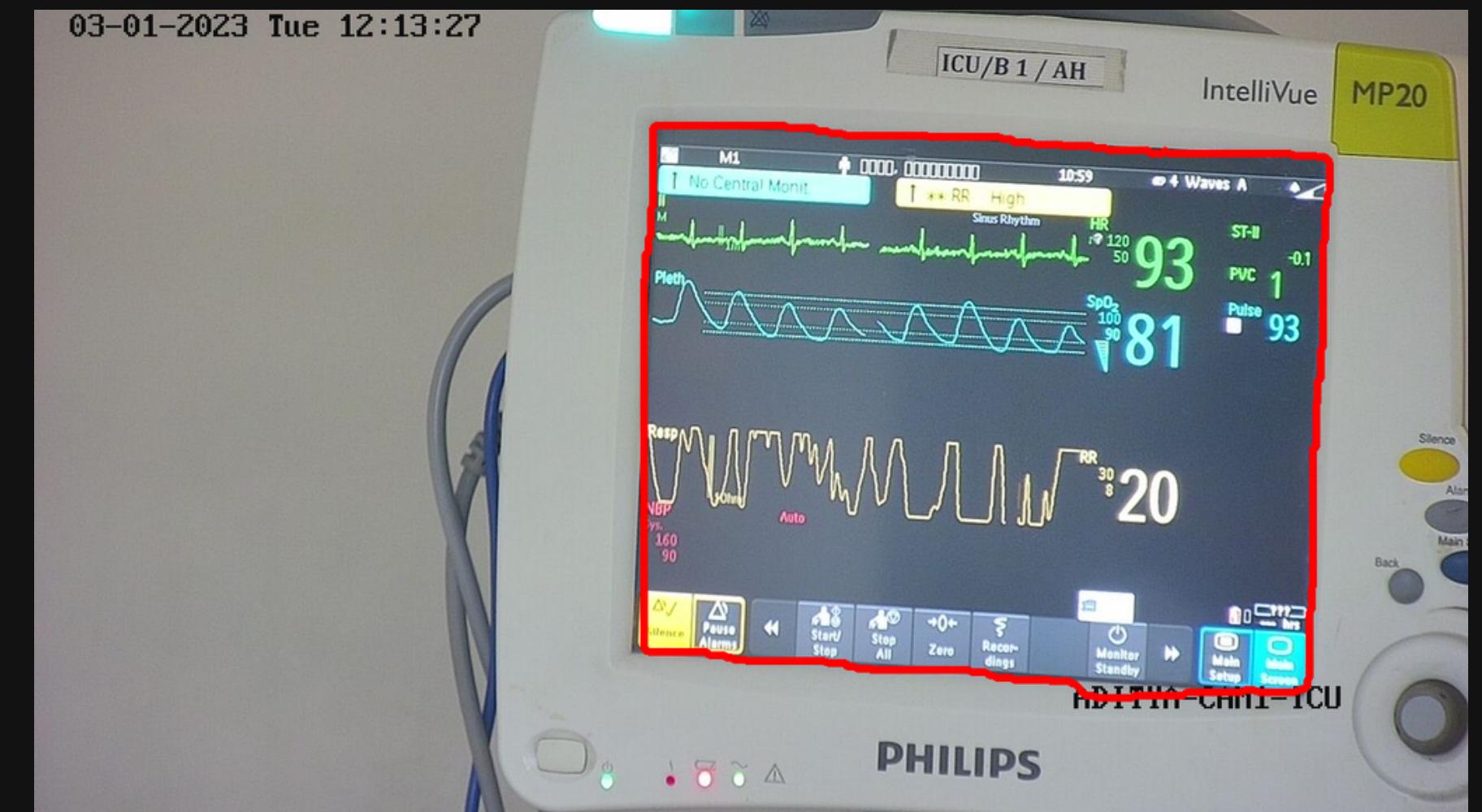
Warp the screen from its corners to erase perspective distortion



Screen Segmentation

Trained YOLO v5

- Low latency
- Learns a general representation
- Gives Screen segmentation



Corner Detection

Used OpenCV

- Created approximate quadrilateral bounding box
- Gave only 4 bounding box corners per image
- Facilitates image warping later

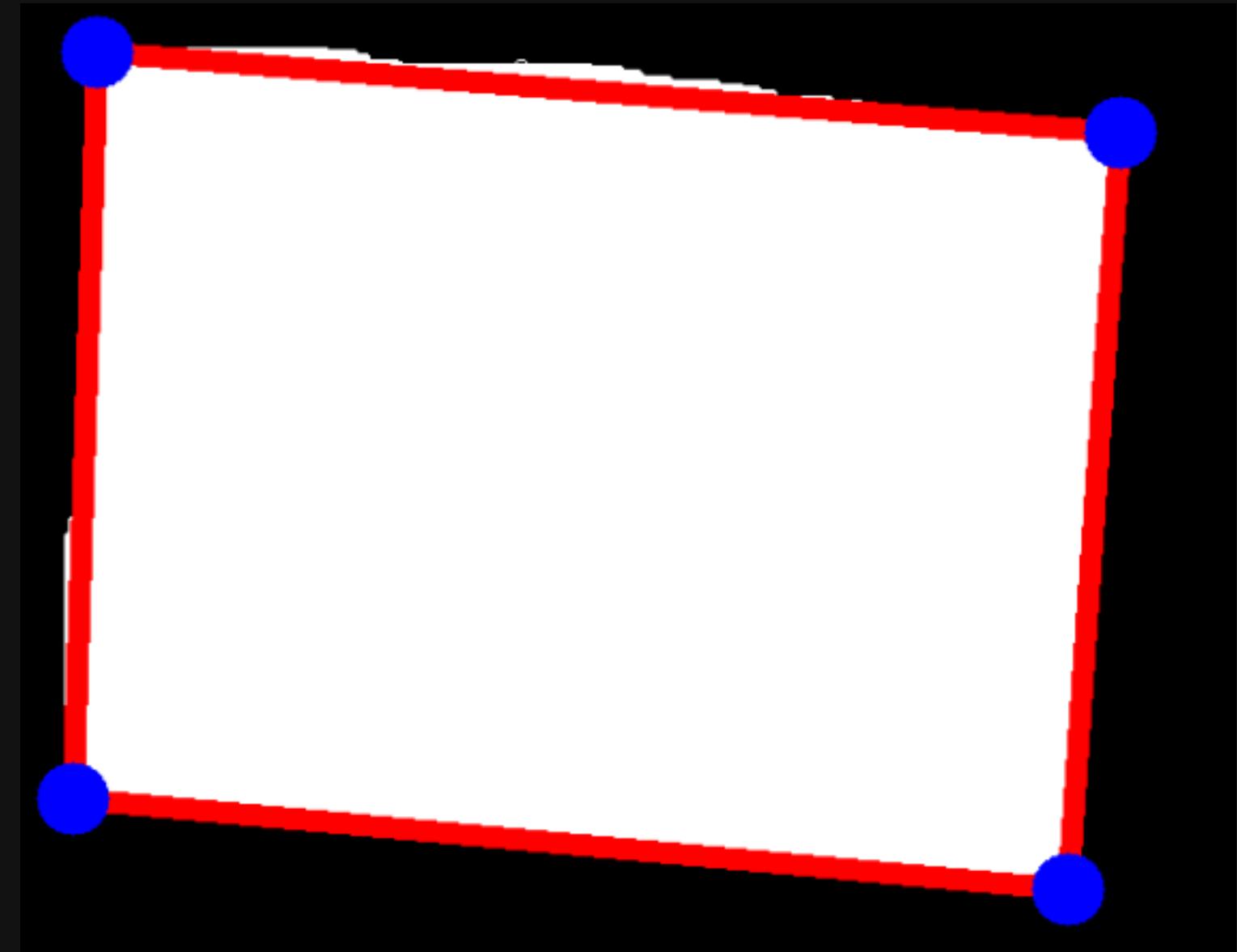
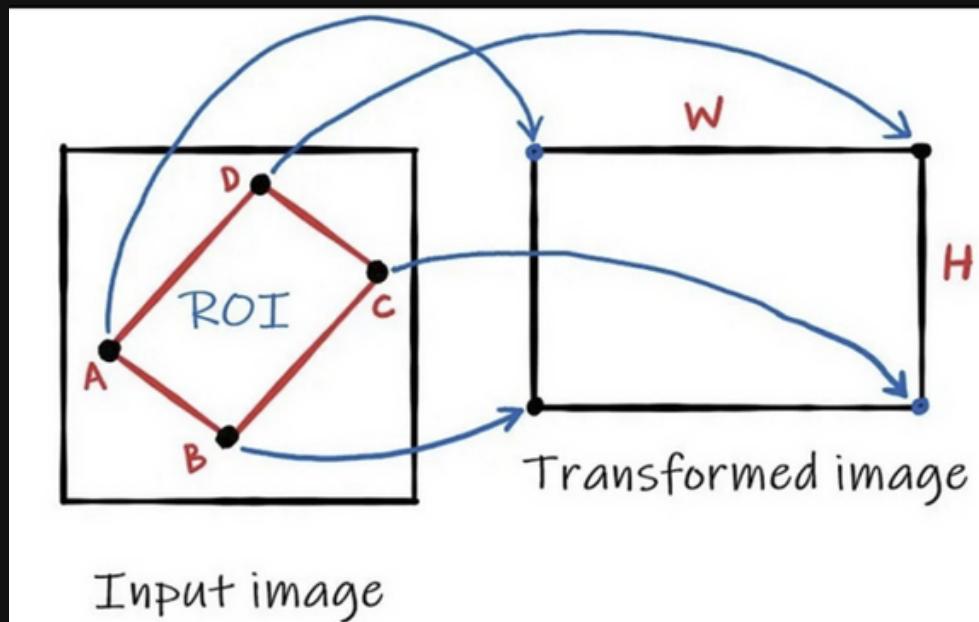
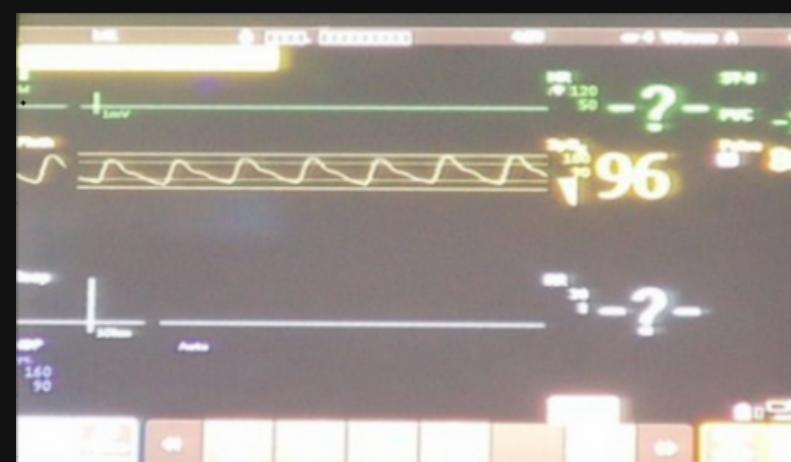


IMAGE WARPING



Perspective Transform

Generate a warping matrix for the segmented monitor using the initial and final corner coordinates



Warping Step

Use the generated image and warp the image into a front facing orientation

VITAL CLASSIFICATION PIPELINES

1 TEMPLATE MATCHING

Custom CNN based approach to find monitor type

2 ERODED IMAGES

Localization of data based on surrounding symbols

3 GRID MATCHING

Localization of data based on surrounding symbols

4 NEIGHBORHOOD MATCHING

Template Independent Pipeline

1 Template Matching



Methods

If variety of monitor layouts is little, and every monitor has a fixed layout

Create templates of available monitor types and learn the position of data

Identify each metric using the color of the text in the display

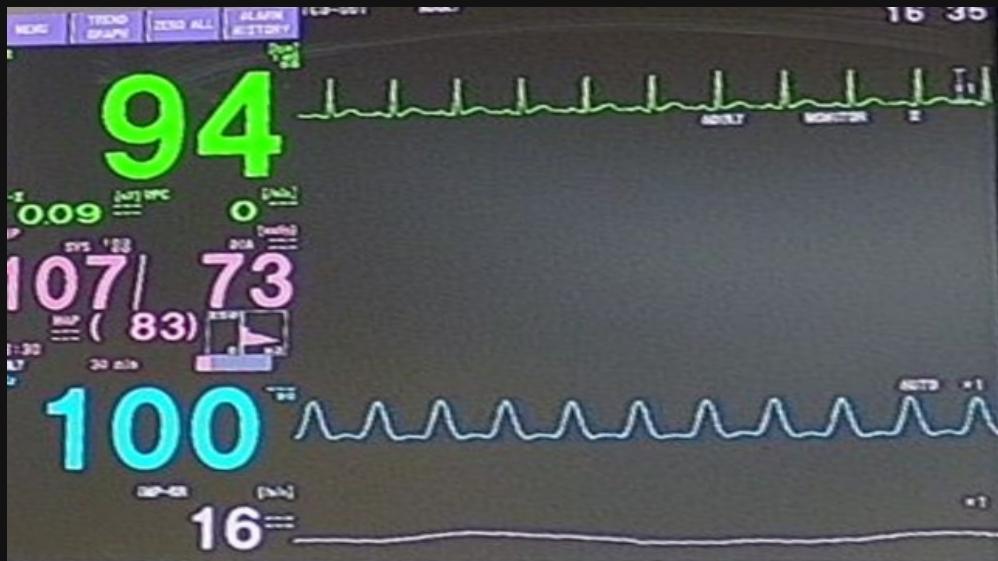
Drawbacks

Huge variety in monitor layouts

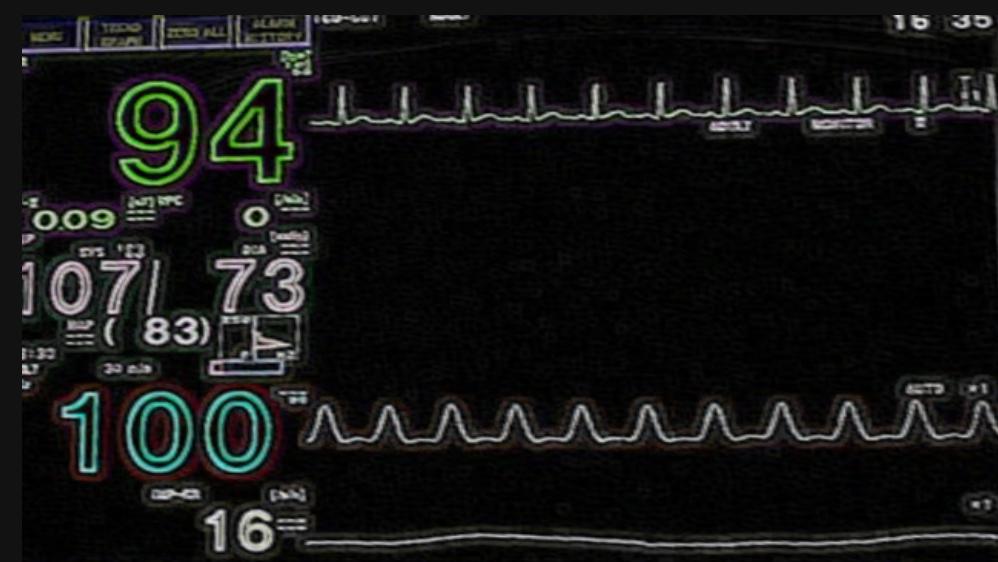
Labels were often swapped even in similar layout patterns

Colors often changed for the same label, and colors were duplicated in many cases

2 ERODED IMAGES PIPELINE



Warped Image



Eroded Image

Bounding Box Detection and
Label Mapping

YOLOv5

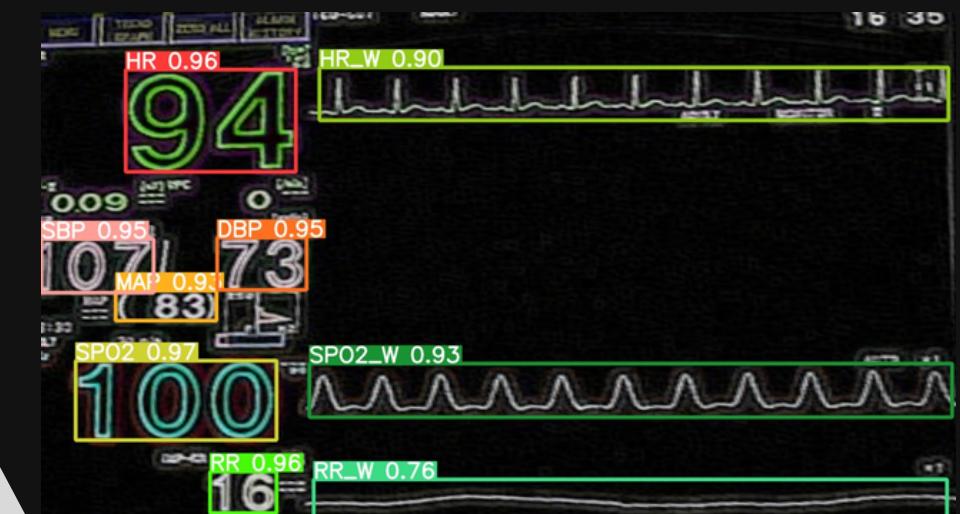
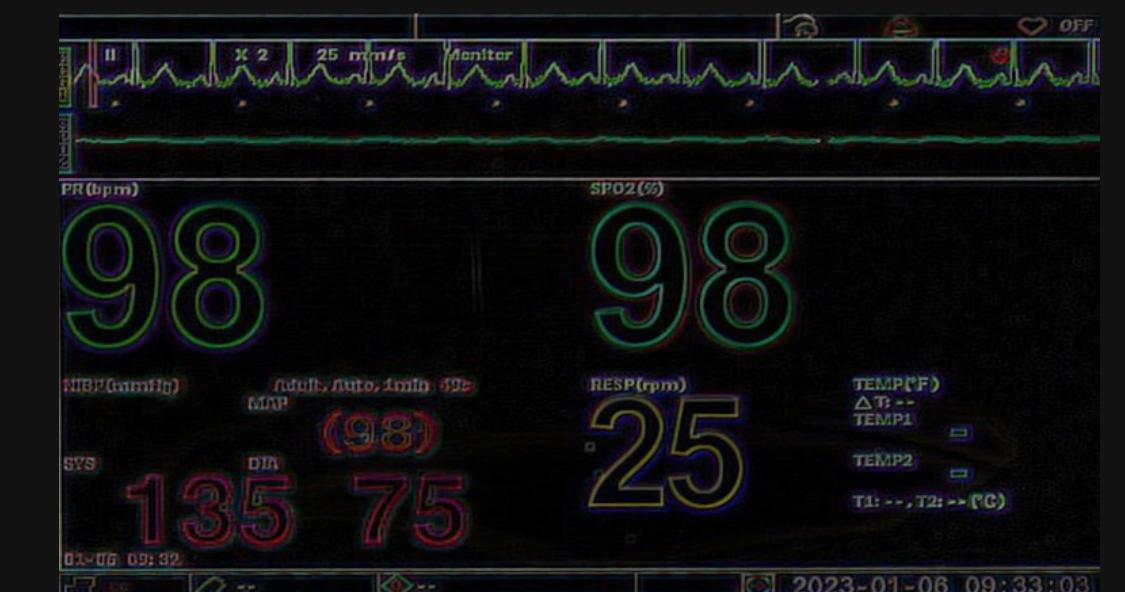
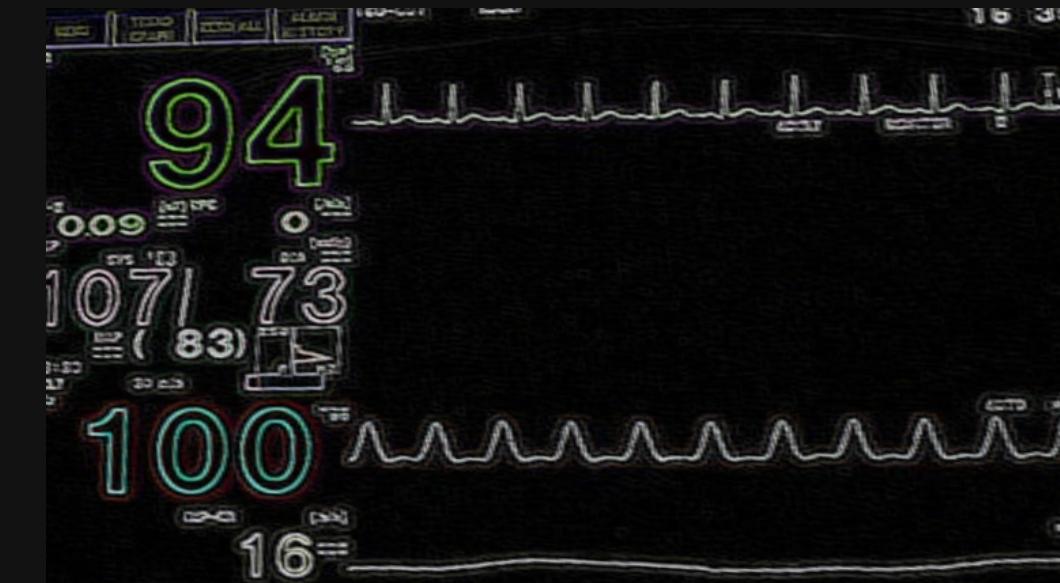
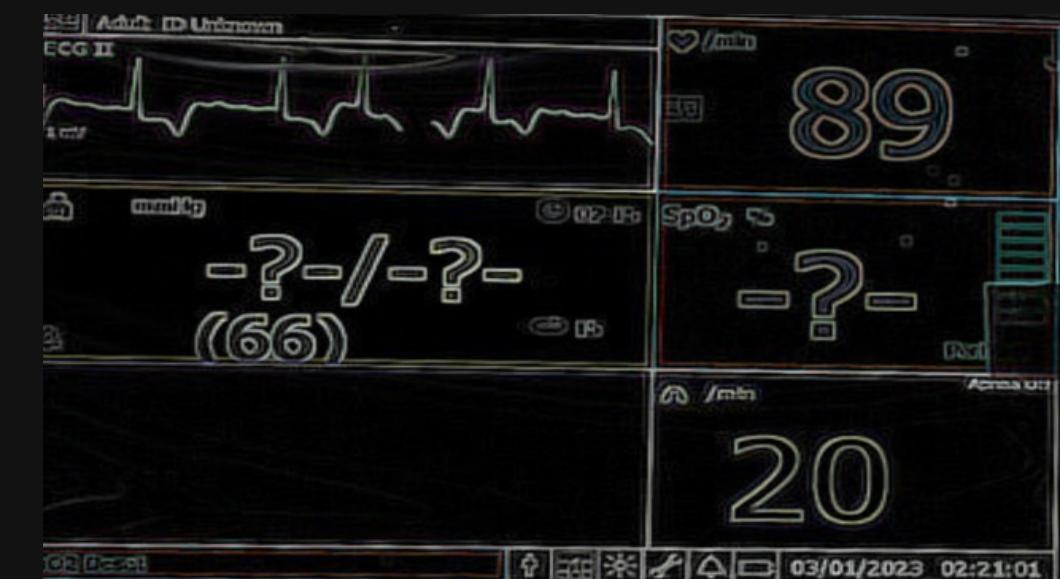


Image Processing Used

Eroded Image Subtraction

No Reflection
No Glare
Background Removed
Cleared out number boundaries



3 GRID PIPELINE



Labelled the whole
grid containing vital
information.

Better YOLO
detection due to
neighbouring
elements

TEXTS
SYMBOLS
Value Range

Procedure

01

Labelled the Whole Grid
containing VITAL INFORMATION

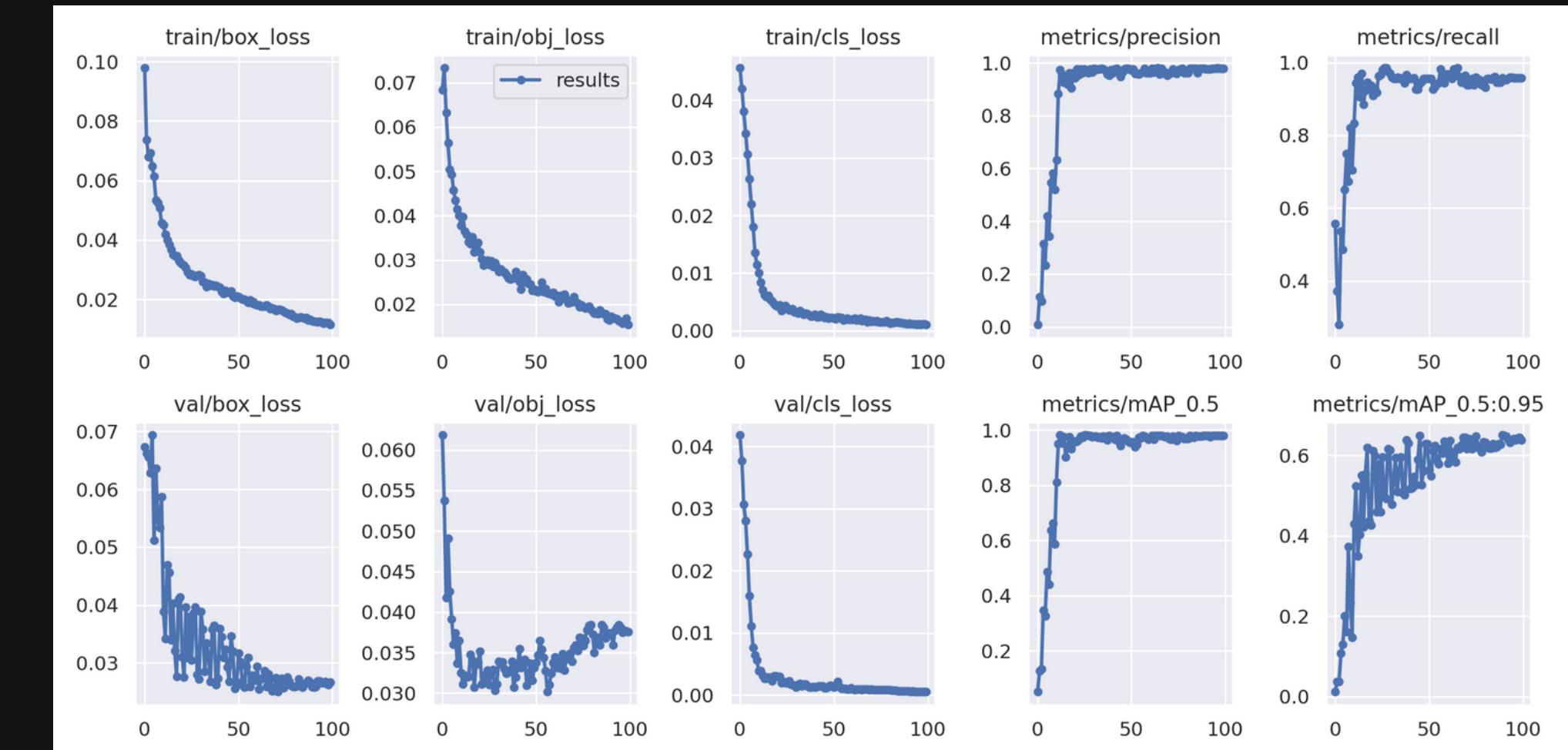
Also includes TEXTS, SYMBOLS and other
relevant information.



02

Trained on YOLOv5

Used various DataAugmentation methods like
Rotation, Grayscale, Shear etc.



DETECTION on Results

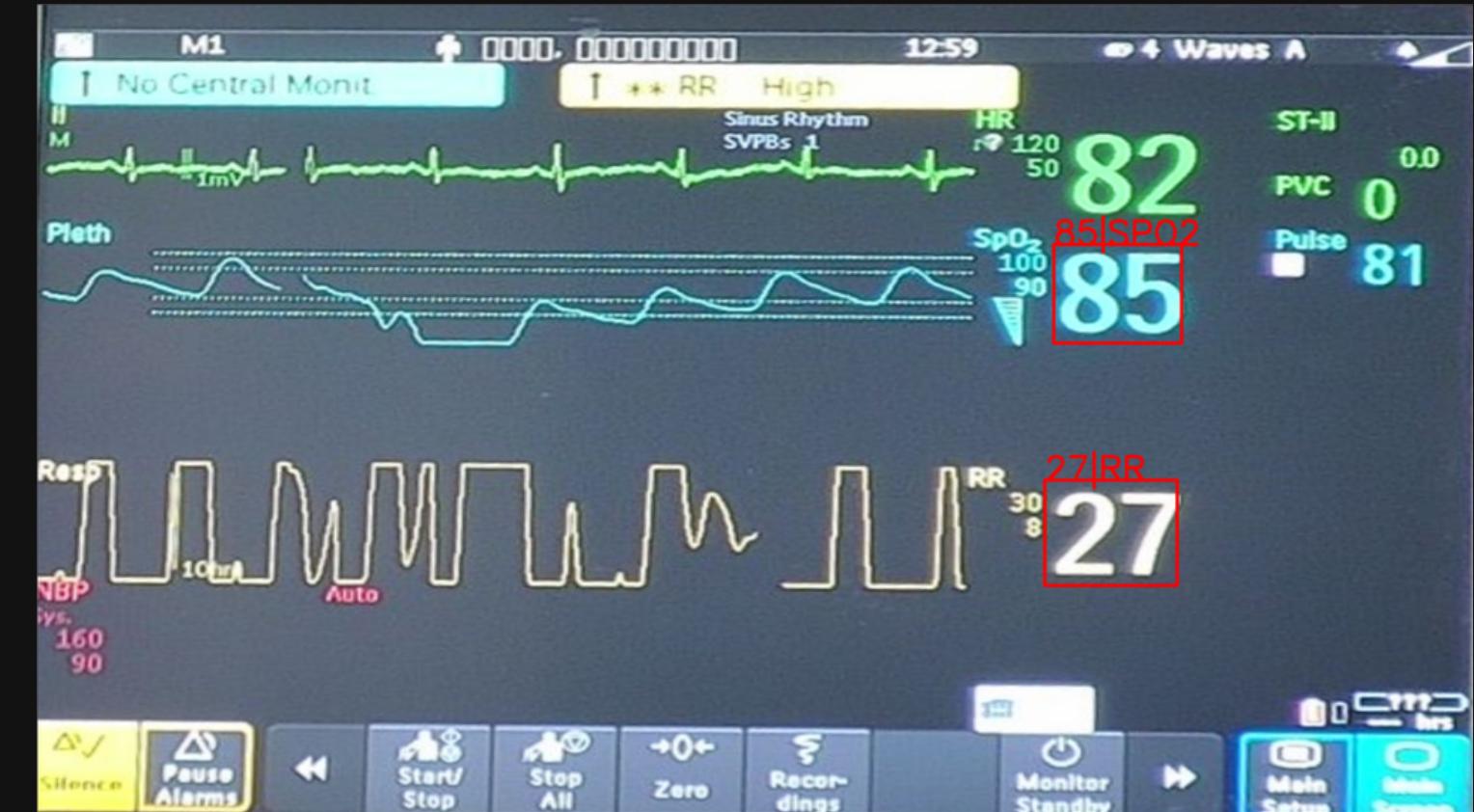
Fast and Accurate Detection of VITAL Classes



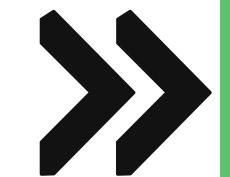
4

NEIGHBORHOOD MATCHING

Template - independent pipeline



TEXT DETECTION



TEXT RECOGNITION



CORRELATION

TEXT DETECTION MODELS

SAST



FCE



Text Recognition Models

NRTR



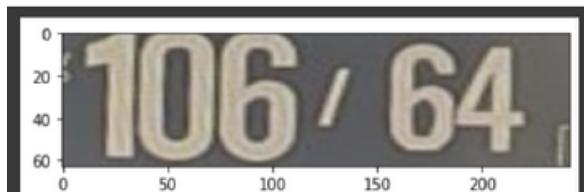
```
[2023/02/07 15:21:58] ppocr INFO: In PP-OCRv3, rec_image_shape parameter defaults to '3, 48, 320', if you are using recognition model with PP-OCRv2 or an older vers  
[2023/02/07 15:21:59] ppocr INFO: Predicts of ./content/smol_test/4_4.jpeg:('eeeeeeeeeeeeeeeeeeeeeee', 0.21478259563446045)
```

Outputs



106/64

SAR



```
[2023/02/07 15:25:16] ppocr INFO: In PP-OCRv3, rec_image_shape parameter defaults to '3, 48, 320', if you are using recognition model with PP-OCRv2 or an older vers  
[2023/02/07 15:25:22] ppocr INFO: Predicts of ./content/smol_test/4_4.jpeg:('106|64-----', 0.9700289368629456)
```

RFL



```
[2023/02/07 01:37:20] ppocr INFO: In PP-OCRv3, rec_image_shape parameter defaults to '3, 48, 320', if you are using recognition model with PP-OCRv2 or an older vers  
[2023/02/07 01:37:20] ppocr INFO: Predicts of ./content/smol_test/4_4.jpeg:('10664', 0.9990275502204895)
```

10664

SVTR



```
[2023/02/07 15:20:29] ppocr INFO: In PP-OCRv3, rec_image_shape parameter defaults to '3, 48, 320', if you are using recognition model with PP-OCRv2 or an older vers  
[2023/02/07 15:20:29] ppocr INFO: Predicts of ./content/smol_test/4_4.jpeg:('10664', 0.9981959462165833)
```

10664

CORRELATION

01

Area-wise Sorted Bounding Boxes

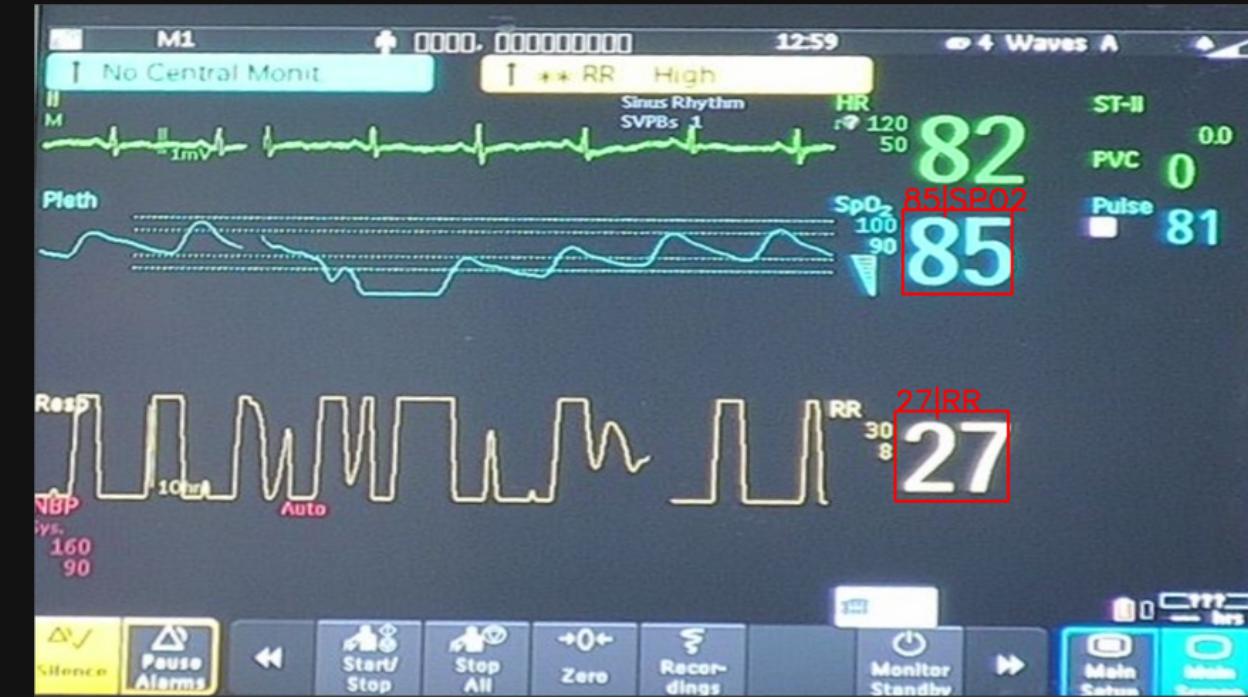
Bigger bounding boxes contain vital data,
smaller ones contain labels



02

Vital Data Labelling

Nearby text labels on the top were read and mapped to the numeric data.



OCR Testing

Some of the OCR libraries we tried

KerasOCR

Very long processing time

TesseractOCR

Requires specific training for
our use case



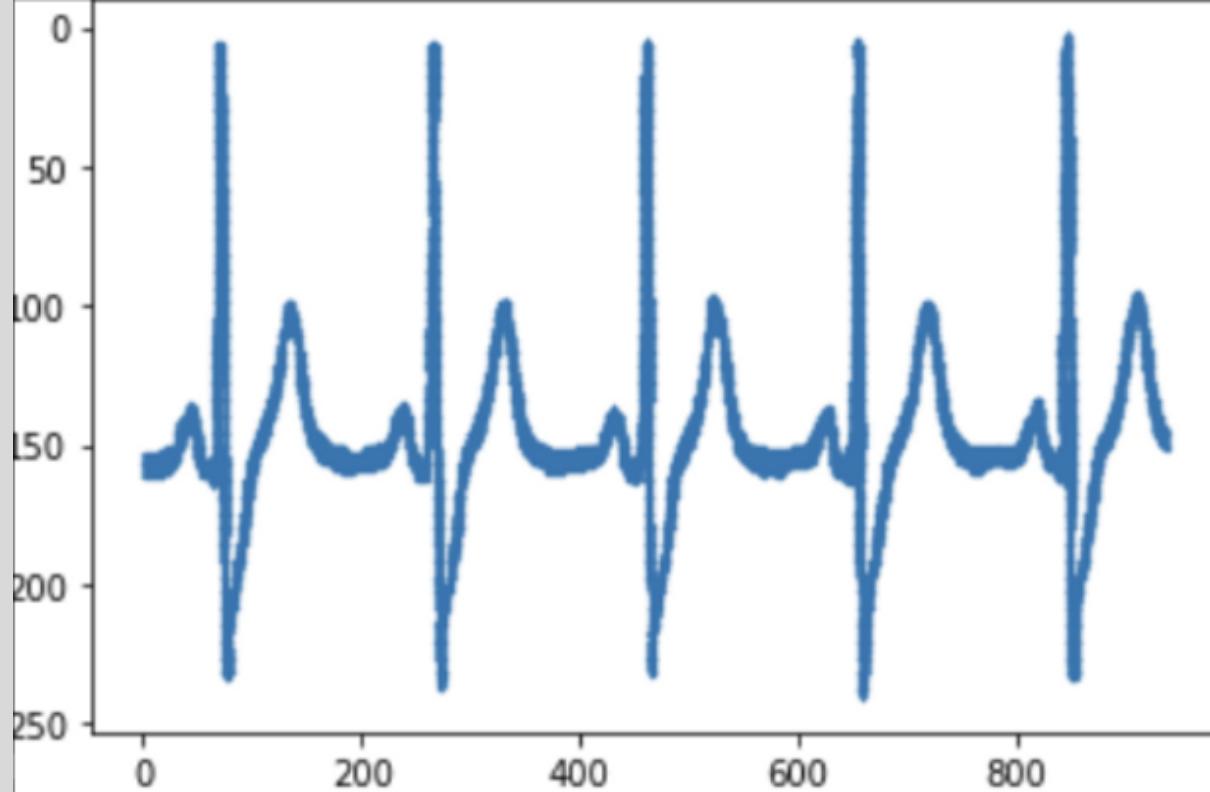
easyOCR

PaddleOCR



Novelties

HR Graph Digitization

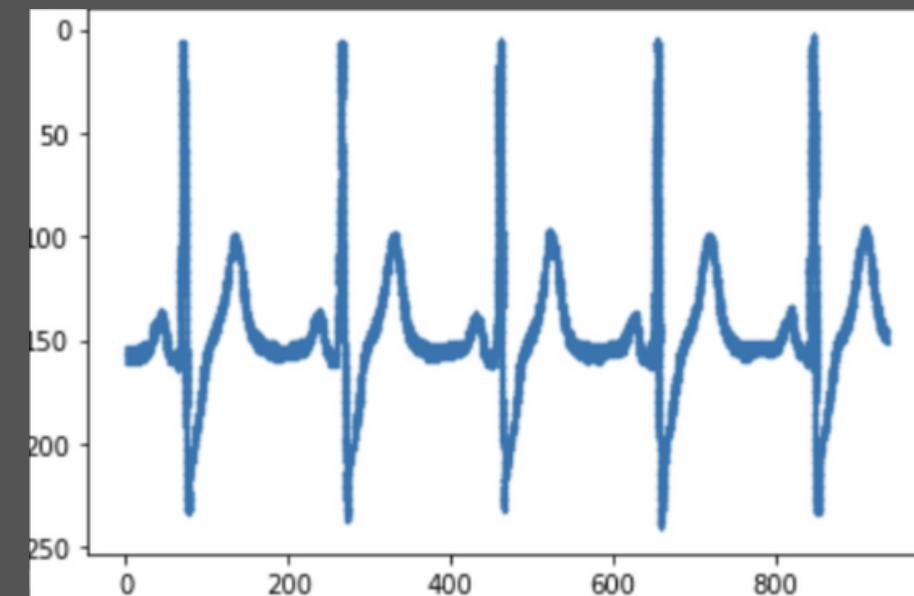


Emergency Light Detection



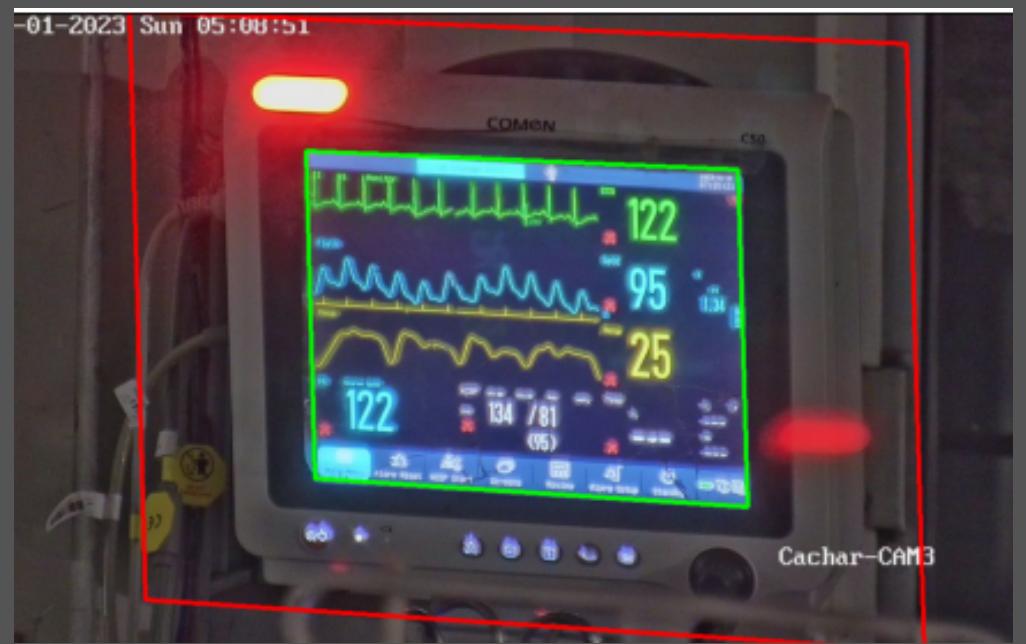
Heart Rate Digitization

- The HR graph is identified and segmented out of the monitor screen
- Image processing is applied to obtain a binarized image
- The binarized image is plotted onto a graph



Emergency Light Detection

- Monitor images are segmented between two bounding boxes
- The image is converted to HSV and red elements are filtered
- Red pixel density in the image determines the presence of flashing lights in the image



FUTURE IMPROVEMENTS

- 1 Improve the inference time by using better or smaller models
- 2 GANs can be used to improve the image quality before detecting
- 3 Deploy a lighter model for EDGE devices so that it works on CCTV footage.
- 4 Fourier transform to decide the pipeline into which the images will be passed on.

THANK YOU

Any Questions?