System Design Workflow for Staff Detection via Name Tag Recognition

# Detection Workflow (Idea for local deployment)

1. *Video Input*
   * Read video file, extract frames at fixed intervals.
2. *Staff Detection*
   * Detect objects <staff, staff\_tag> using a pre-trained YOLO model.
   * If any staff has a name tag, flag frame as containing staff.
3. *Coordinate Extraction*
   * Record bounding box as coordinates of the staff’s bounding box.
   * Save staff presence frame.
4. *Output and Reporting*
   * To run with a different test video/ images.

# Step 1: Modelling

[Colab Notebook](https://colab.research.google.com/drive/1dIbm6Ld0lTjNXL67gQj5yiMSF4YMX_ye#scrollTo=1sUfcA8ZgR2t) for training model -> save my\_model.zip in locally

## Data exploration

Video (sample.mp4) serves as the raw input data for analysis. The video is then decomposed into individual image frames at consistent time intervals (25 frames per second). This sampling rate is selected to maintain a balance between detection accuracy and computational efficiency. Each extracted frame is prepared for further processing in the subsequent stages. Following by Training, Validation and Testing metrics.

|  |  |  |  |
| --- | --- | --- | --- |
| **Original dataset** | **Total data** | **With Objects** | **W/O Objects (Background)** |
| Train | 879 | 708 | 171 (20%) |
| Validation | 36 | 35 | 1 |
| Test | 37 | 34 | 3 |

[Roboflow labelled dataset](https://universe.roboflow.com/staff-detection-snwgo/staff-detection-njhrv/dataset/5)

## Training Metrics

|  |  |
| --- | --- |
| **Column** | **Meaning** |
| box\_loss | Bounding box regression loss — lower = better localization |
| cls\_loss | Classification loss |
| dfl\_loss | Distribution Focal Loss |
| Images | Number of validation images evaluated (57) |
| Instances | Total objects in the validation set (98 bounding boxes) |
| Box(P) | Precision = TP / (TP + FP) → % predictions are correct |
| R | Recall = TP / (TP + FN) → % ground truth was detected |
| mAP50 | Mean Average Precision at IoU ≥ 0.50 |
| mAP50-95 | Mean Average Precision averaged from IoU 0.50 to 0.95 |

Training configuration: epochs=60 imgsz=640. Below is the training results.

A blue squares with white text

AI-generated content may be incorrect.

A graph of a graph

AI-generated content may be incorrect.A diagram of a graph

AI-generated content may be incorrect.A close-up of a graph

AI-generated content may be incorrect.A group of graphs showing the number of data

AI-generated content may be incorrect.A group of people in a room

AI-generated content may be incorrect.