Person and PPE Detection Project Report

Approach

1. Dataset Preparation

- Implemented conversion from PascalVOC to YOLO format
- Created separate datasets for person detection and PPE detection
- Balanced classes by filtering and sampling

2. Two-Stage Detection System

1. **Person Detection**:
    - Used YOLOv8 model trained on full images
    - Focus on accurate person detection as foundation
    - Implemented coordinate system for cropping
2. **PPE Detection**:
    - Trained on cropped person regions
    - Selected 5 key PPE classes:
        - hard-hat
        - gloves
        - boots
        - vest
        - ppe-suit
    - Implemented coordinate transformation for mapping detections back to original images

3. Implementation Logic

1. **Annotation Conversion**:

```
def convert_coordinates(box, img_width, img_height):
    # Convert from PascalVOC (xmin, ymin, xmax, ymax) to YOLO format
    x_center = ((box[0] + box[2]) / 2) / img_width
    y_center = ((box[1] + box[3]) / 2) / img_height
    width = (box[2] - box[0]) / img_width
    height = (box[3] - box[1]) / img_height
    return [x_center, y_center, width, height]
```

2. **PPE Detection on Cropped Images**:

```
def adjust_ppe_coordinates(ppe_box, person_box, orig_dims):
    # Transform PPE coordinates relative to person crop
    x1 = (ppe_box[0] - person_box[0]) / person_box[2]
    y1 = (ppe_box[1] - person_box[1]) / person_box[3]
    x2 = (ppe_box[2] - person_box[0]) / person_box[2]
    y2 = (ppe_box[3] - person_box[1]) / person_box[3]
    return [x1, y1, x2, y2]
```

Evaluation Metrics

Person Detection Model

- Trained on full images
- Focus on high recall to ensure no persons are missed

PPE Detection Model

Performance metrics:

- Overall: mAP50 = 0.334, mAP50-95 = 0.199
- hard-hat: mAP50 = 0.821 (Best performing)
- gloves: mAP50 = 0.088
- boots: mAP50 = 0.342
- vest: mAP50 = 0.073
- ppe-suit: mAP50 = 0.345

Learning Outcomes

1. **Two-Stage Detection Benefits**:
    - Improved accuracy for small PPE items
    - Better handling of scale variations

- Reduced false positives

2. **Class Performance Analysis**:

   - Hard hats show best detection due to distinctive shape
   - Gloves and vests show lower performance due to size and variation
   - PPE suits and boots show moderate performance

3. **Technical Insights**:

   - Coordinate transformation crucial for accurate detection
   - Class balancing important for model stability
   - OpenCV integration for visualization