“Lines Extraction from ID Documents” Task report

Prepared by : Ahmed Ezzat

Submission Date: 15/1/2024

Introduction

The goal of this project is to develop a **Lines Extraction System** that processes scanned ID documents to extract individual lines of text as cropped images. The solution should handle varying orientations of IDs and segment the document into individual text lines.

Dataset

I have used two different datasets for the solution:

1. ID detection Dataset (Link)
    - This dataset contains 1400 ID images in different rotations.
2. ID parts detection Dataset (Link)
    - This dataset contains 173 ID images with 965 annotations for different ID parts like Address, Gender, Religion…etc.

Methodology

Libraries used (Ultralytics, OpenCV, etc.) for each part:

1. ID Detection:
    - YOLOv8 model for ID detection.
2. Orientation handling:
    - Using detected class to rotate the ID correctly.
    - Rotation correction with OpenCV library.
3. Line segmentation:
    - YOLOv8 model for ID parts detection and extraction.

Results

1. Performance of the ID Card Detection Model

The YOLOv8 model for detecting ID cards and their orientations was trained and validated on a dataset of Egyptian IDs. The results are summarized as follows:

- **Classes Detected**:
    - id (non-rotated IDs)
    - id_rotated
    - id_rotated_down
    - id_rotated_right
    - no (non-ID objects)
- **Validation Metrics**:
    - **Precision (P)**: The model achieved an overall precision of 0.869, indicating high accuracy in identifying ID cards across all orientations.
    - **Recall (R)**: The overall recall was 0.889, ensuring that most ID cards were detected in the dataset.
    - **mAP@0.5**: The mean Average Precision at IoU threshold 0.5 across all classes was 0.91.

- **Class-Specific Performance**:
  - id achieved the highest precision (0.917) and mAP@0.5 (0.953), reflecting accurate detection of non-rotated IDs.
  - Rotated IDs (id_rotated, id_rotated_down, id_rotated_right) showed slightly lower precision and recall due to the additional complexity of orientation.

---

2. Performance of the ID Parts Detection Model

The YOLOv8 model for detecting various parts of the ID was validated on a separate dataset, with the following results:

- **Classes Detected**:
  - Address, Birth, Factory, Gender, HusName, ID_B, ID_F, Name, Occup, Rel, Start, Status, and end.
- **Validation Metrics**:
  - **Precision (P)**: The model achieved a high precision of 0.929, reflecting its capability to identify ID parts accurately.
  - **Recall (R)**: The recall was 0.973, demonstrating the model's ability to detect most ID parts.
  - **mAP@0.5**: The mean Average Precision at IoU threshold 0.5 was 0.964.
- **Class-Specific Performance**:
  - Parts such as HusName and ID_F achieved near-perfect precision and recall, with mAP@0.5 of 1.0.

Summary

1. **ID Card Detection**:
   - The model is robust, with high precision and recall for detecting ID cards and their orientations.
   - Non-rotated IDs achieved the best detection performance.
2. **ID Parts Detection**:
   - The model effectively detects ID parts with high precision and recall.
   - Classes with more training instances achieved better performance, highlighting the need for balanced datasets for underrepresented classes.

```
val: Scanning /kaggle/working/zkzk-zkzk-4/valid/labels.cache... 395 images, 6 ba
```

Class	Images	Instances	Box(P	R	mAP50	m
all	395	491	0.869	0.889	0.91	0.798
id	62	63	0.917	0.937	0.953	0.902
id_rotated	71	73	0.883	0.945	0.942	0.854
id_rotated_down	59	59	0.928	0.966	0.966	0.887
id_rotated_right	72	73	0.895	0.863	0.925	0.827
no	163	223	0.719	0.734	0.762	0.52

```
Speed: 0.9ms preprocess, 36.4ms inference, 0.0ms loss, 1.3ms postprocess per image
val: Scanning /kaggle/working/CDC_OCR-2-1/valid/labels.cache... 18 images, 0 bac
```

Class	Images	Instances	Box(P	R	mAP50	m
all	18	99	0.93	0.976	0.963	0.706
Address	11	11	0.892	1	0.912	0.763
Birth	11	12	0.894	0.917	0.928	0.644
Factory	10	10	0.734	0.83	0.817	0.457
Gender	6	6	0.96	1	0.995	0.64
HusName	2	2	0.995	1	0.995	0.801
ID_B	6	6	0.945	1	0.995	0.727
ID_F	11	11	0.895	1	0.965	0.632
Name	11	11	0.892	1	0.942	0.782
Occup	6	6	0.956	1	0.995	0.902
Rel	6	6	0.967	1	0.995	0.724
Start	6	6	0.982	1	0.995	0.776
Status	6	6	1	0.948	0.995	0.576
end	6	6	0.981	1	0.995	0.76

```
Speed: 5.1ms preprocess, 33.1ms inference, 0.0ms loss, 10.2ms postprocess per image
```