

Face Detection, Recognition & Tracking — Evaluation Report

1. Goal of the Project

The goal of this project was to implement and evaluate a face detection, tracking, and recognition system on video data. The main requirement was to achieve at least 80% Recall (TPR).

2. Methodology

Several detectors were tested, including Haar, LBP, and DLib. The best-performing configuration was a combined DLib frontal detector with Haar profile detection. Tracking was performed using the KCF tracker.

```
== Сравнение всех детекторов на person_01_v1 ==

haar:
    TP (True Positive): 35
    FP (False Positive): 447
    FN (False Negative): 6
    TPR (Recall):      85.3659%

haar_profile:
    TP (True Positive): 36
    FP (False Positive): 482
    FN (False Negative): 5
    TPR (Recall):      87.8049%

lbp:
    TP (True Positive): 0
    FP (False Positive): 0
    FN (False Negative): 41
    TPR (Recall):       0%

lbp_profile:
    TP (True Positive): 0
    FP (False Positive): 0
    FN (False Negative): 41
    TPR (Recall):       0%

dlib:
    TP (True Positive): 13
    FP (False Positive): 1
    FN (False Negative): 28
    TPR (Recall):       31.7073%

dlib_profile:
    TP (True Positive): 36
    FP (False Positive): 483
    FN (False Negative): 5
    TPR (Recall):       87.8049%
```

Figure 1. Detector comparison on person_01_v1

3. Overall Results — Known Persons

20 videos were processed.

Overall Recall (TPR): 88.01%

Average Recall per video: 83.45%

Precision: ~9.3%

True Positives: 1880

False Positives: 18314

False Negatives: 256

SUMMARY: person_	
Обработано видео:	20
Overall TPR (Recall):	88.01%
Average TPR (средний по видео):	83.45%
Overall Precision:	9.30%
Total TP (True Positive):	1880
Total FP (False Positive):	18314
Total FN (False Negative):	256
Детальные результаты:	
✓ person_01_v1	TPR: 87.8049%, TP: 36, FP: 483, FN: 5
✓ person_01_v2	TPR: 99.4286%, TP: 174, FP: 288, FN: 1
✓ person_02_v1	TPR: 63.9098%, TP: 85, FP: 2460, FN: 48
✓ person_02_v2	TPR: 90.625%, TP: 29, FP: 87, FN: 3
✓ person_03_v1	TPR: 84.2105%, TP: 48, FP: 607, FN: 9
✓ person_03_v2	TPR: 86.4035%, TP: 197, FP: 2060, FN: 31
✓ person_04_v1	TPR: 89.3617%, TP: 42, FP: 573, FN: 5
✓ person_04_v2	TPR: 84.2593%, TP: 91, FP: 1195, FN: 17
✓ person_05_v1	TPR: 67.9245%, TP: 36, FP: 646, FN: 17
✓ person_05_v2	TPR: 98.3607%, TP: 240, FP: 2689, FN: 4
✓ person_06_v1	TPR: 84.9057%, TP: 45, FP: 835, FN: 8
✓ person_06_v2	TPR: 93.8776%, TP: 46, FP: 910, FN: 3
✓ person_07_v1	TPR: 62.5%, TP: 25, FP: 298, FN: 15
✓ person_07_v2	TPR: 84.8739%, TP: 101, FP: 617, FN: 18
✓ person_08_v1	TPR: 99.3576%, TP: 464, FP: 1530, FN: 3
✓ person_08_v2	TPR: 79.5918%, TP: 39, FP: 560, FN: 10
✓ person_09_v1	TPR: 73.8562%, TP: 113, FP: 1174, FN: 40
✓ person_09_v2	TPR: 86.3636%, TP: 19, FP: 172, FN: 3
✓ person_10_v1	TPR: 76.4706%, TP: 26, FP: 723, FN: 8
✓ person_10_v2	TPR: 75%, TP: 24, FP: 407, FN: 8

Figure 2. Summary results for person videos

4. Overall Results — Unknown Persons

10 videos were processed.

Overall Recall (TPR): 89.10%

Average Recall per video: ~91%

True Positives: 1872

False Positives: 39140

False Negatives: 229

SUMMARY: unknown_	
Обработано видео:	10
Overall TPR (Recall):	89.10%
Average TPR (средний по видео):	90.95%
Overall Precision:	4.56%
Total TP (True Positive):	1872
Total FP (False Positive):	39140
Total FN (False Negative):	229
Детальные результаты:	
✓ unknown_01	TPR: 100%, TP: 1, FP: 0, FN: 0
✓ unknown_02	TPR: 93.9252%, TP: 201, FP: 1091, FN: 13
✓ unknown_03	TPR: 92.3077%, TP: 84, FP: 266, FN: 7
✓ unknown_04	TPR: 96.732%, TP: 148, FP: 3020, FN: 5
✓ unknown_05	TPR: 82.2055%, TP: 328, FP: 8502, FN: 71
✓ unknown_06	TPR: 81.0427%, TP: 171, FP: 4697, FN: 40
✓ unknown_07	TPR: 91.0072%, TP: 253, FP: 6474, FN: 25
✓ unknown_08	TPR: 94.2761%, TP: 280, FP: 6834, FN: 17
✓ unknown_09	TPR: 89.6774%, TP: 139, FP: 3476, FN: 16
✓ unknown_10	TPR: 88.4106%, TP: 267, FP: 4780, FN: 35

Figure 3. Summary results for unknown videos

5. Conclusion

This project successfully implemented and evaluated a face detection, tracking, and recognition system on video data.

The experimental results demonstrate that the proposed system meets the main project requirement, achieving an overall recall above the required 80% threshold for both known and unknown subjects. The evaluation confirms that combining multiple detectors (DLib frontal and Haar profile) with tracking significantly improves detection robustness across different face orientations and video conditions.

In addition to quantitative metrics, qualitative results further support the effectiveness of the system. As shown in Figure 4, Figure 5 and Figure 6- the detector is able to accurately localize faces in real video frames, providing stable bounding boxes suitable for subsequent tracking and recognition stages.

Overall, the system demonstrates reliable performance and provides a solid baseline for further improvements, such as reducing false positives and enhancing recognition accuracy in open-set



Figure 4. Person_01_v2_frame_00240_annotation

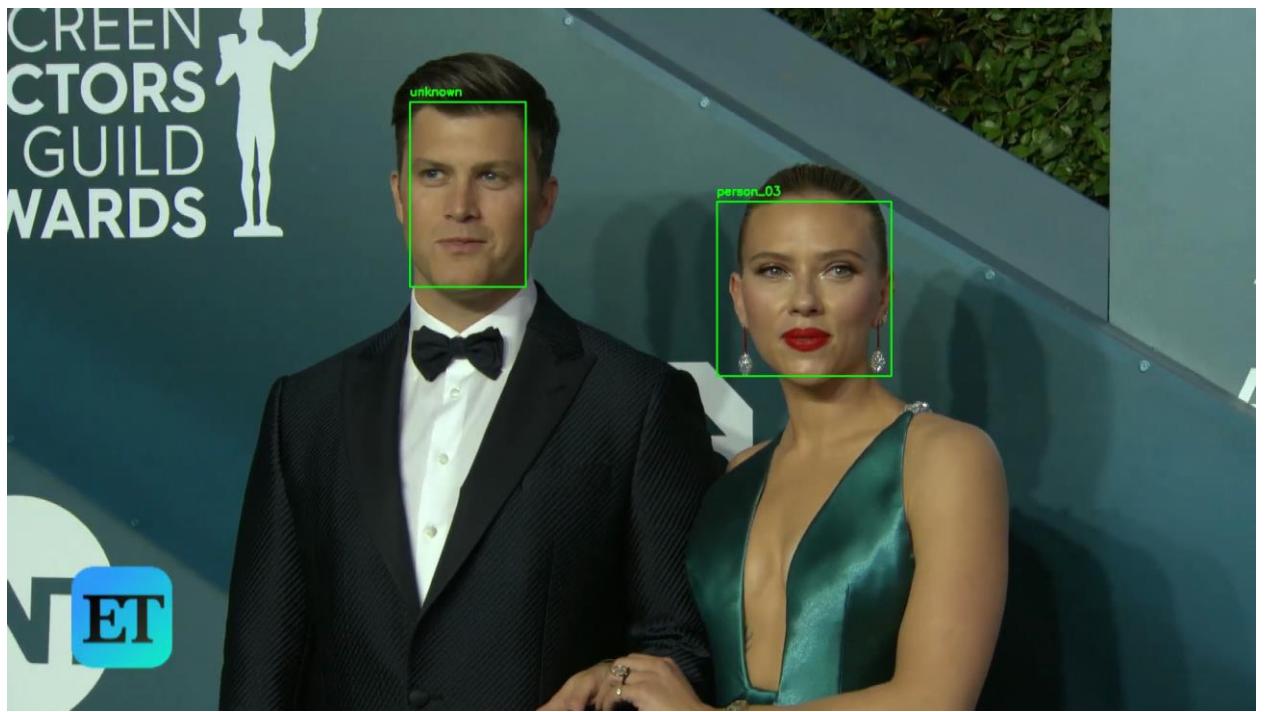


Figure 5. Person_03_v2_frame_01045_annotation

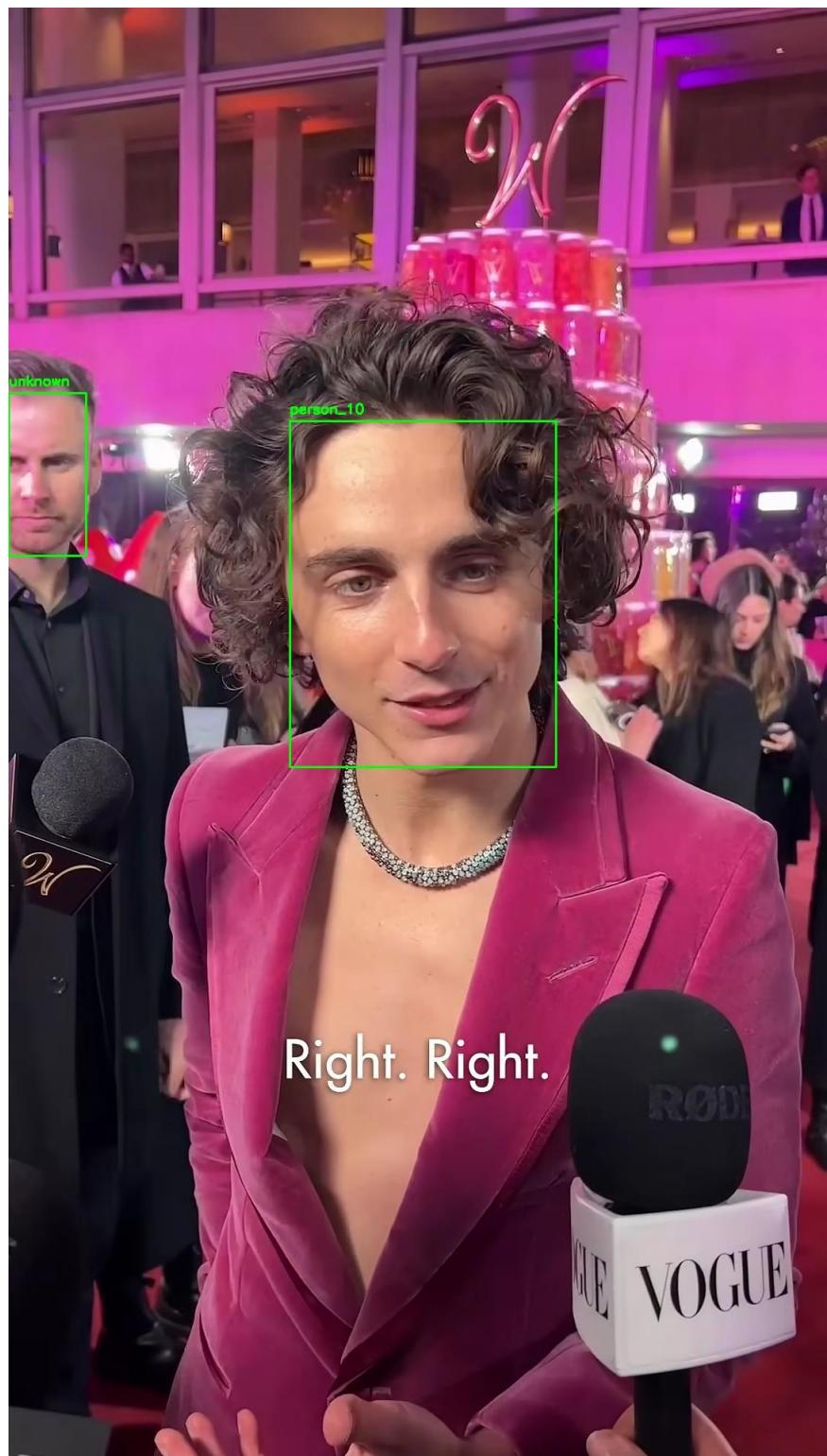


Figure 6. Person_10_v2_frame_00455_annotation