Comparative Analysis of Facial Landmark Detection Models

Dlib vs. MediaPipe

June 5, 2025

Introduction

Purpose of the Analysis:

- Compare performance of two facial landmark detection models
- Evaluate speed, accuracy, and usability

Models Evaluated:

- Dlib: 68-point facial landmark detector
- MediaPipe: Face mesh with 468-point landmarks

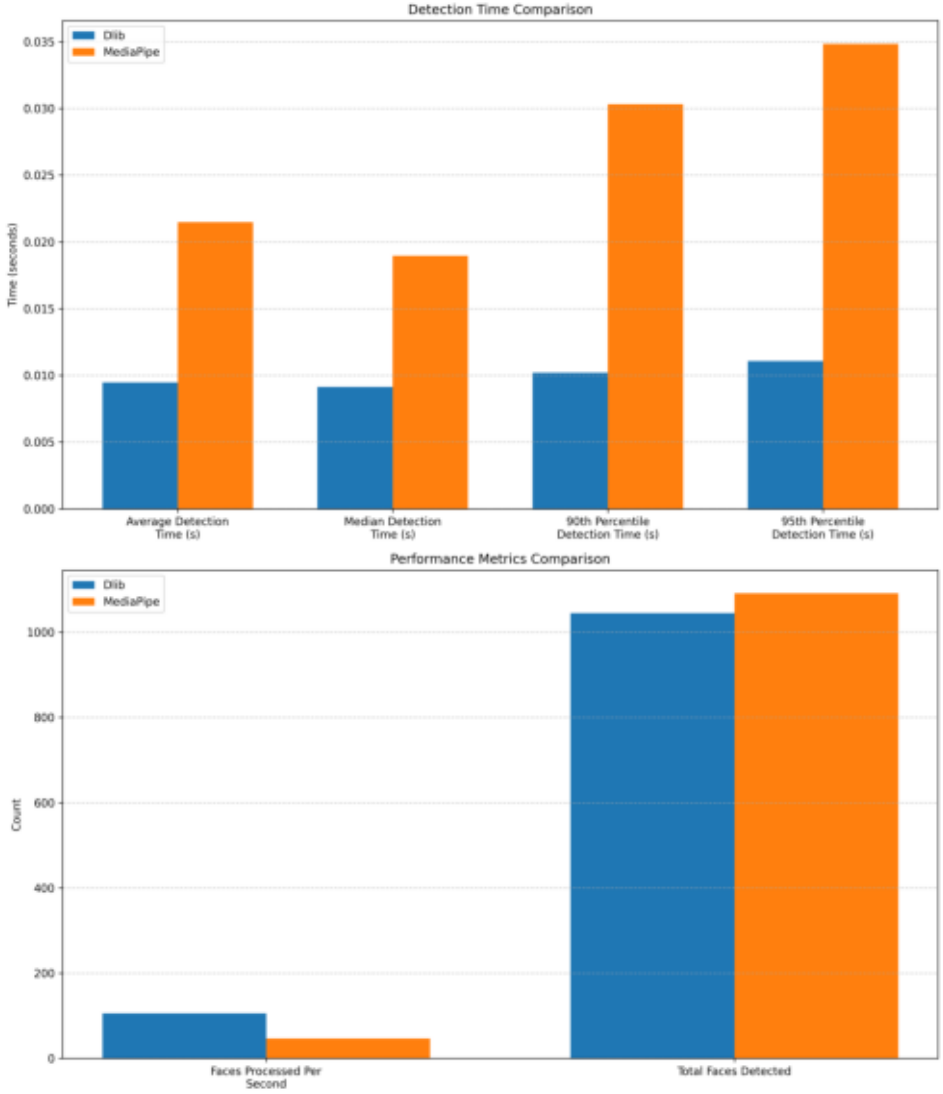
Testing Dataset:

- 1000 facial images with various conditions

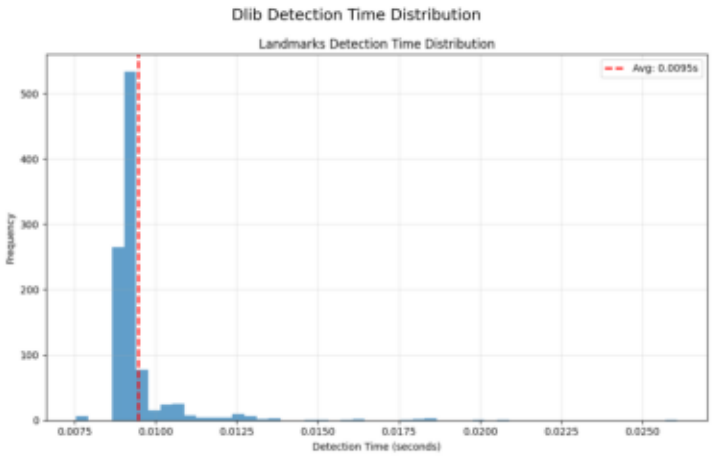
Key Performance Metrics

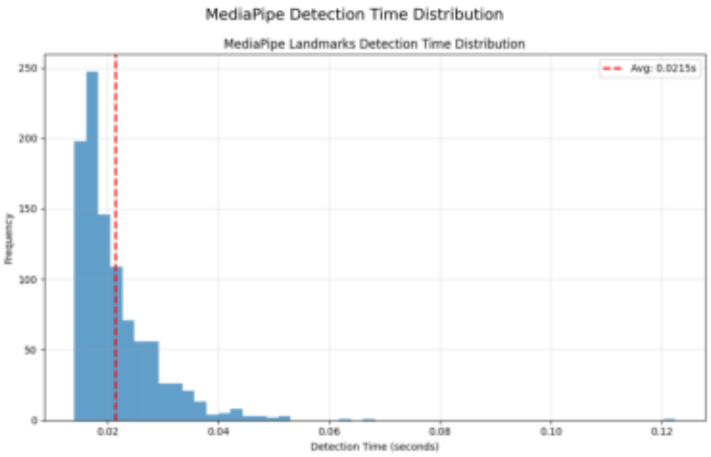
	Metric	Dlib
on Time (s)	0.0095	0.0215
cond	105.69	46.56
Detected	1045	1091
Faces	994	1000

Performance Comparison

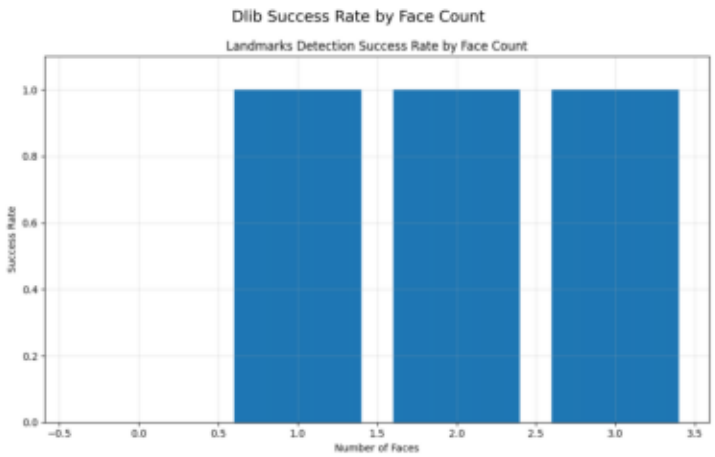


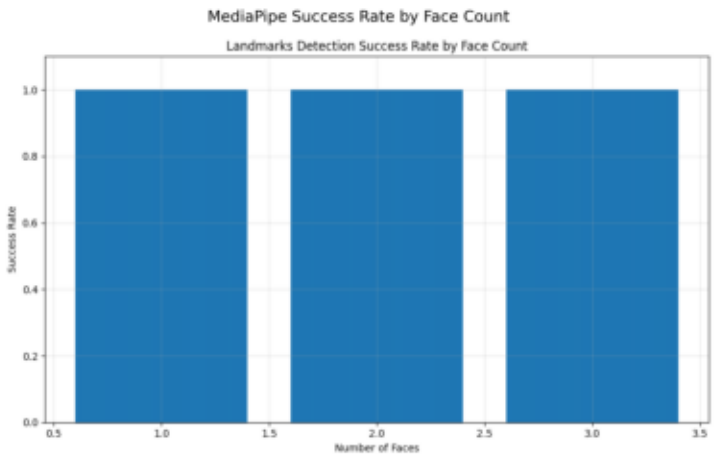
Detection Time Analysis





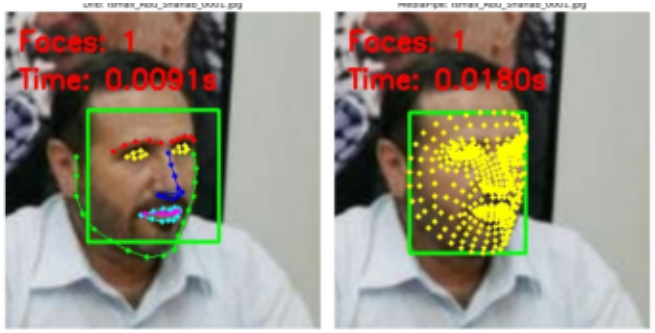
Success Rate Analysis



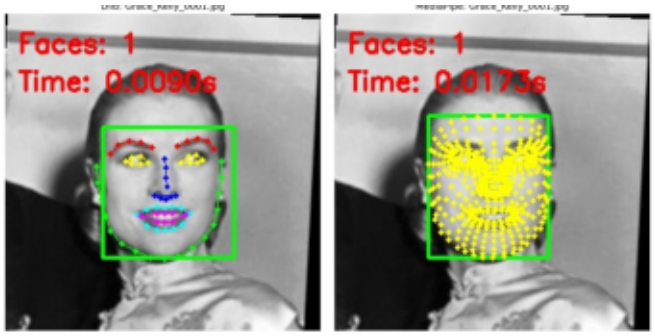


Sample Visualizations: Dlib vs MediaPipe

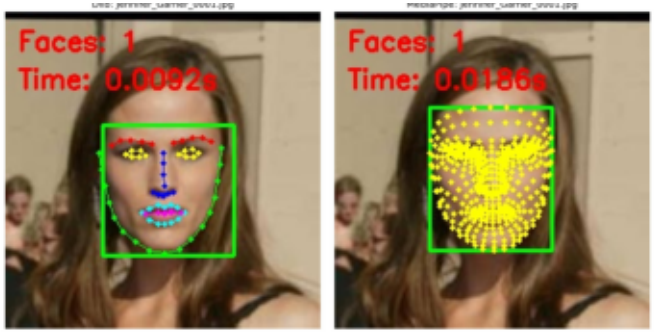
Sample 1: 000_Ismail_Abu_Shanab_0001



Sample 2: 001_Grace_Kelly_0001



Sample 3: 002_Jennifer_Garner_0001



Key Findings

1. Speed vs. Detail Trade-off:

- Dlib is ~2.3x faster than MediaPipe
- MediaPipe provides ~6.9x more landmark points

2. Detection Accuracy:

- MediaPipe detected 4.4% more faces overall
- MediaPipe found faces in all test images

3. Use Case Recommendations:

- Dlib for speed-critical applications
- MediaPipe for precision-critical applications

Conclusion and Recommendations

Summary:

- Both models are effective for facial landmark detection
- Key trade-off is between speed and detail

Recommended Use Cases:

- Dlib: Mobile applications, real-time systems, resource-constrained environm
- MediaPipe: Detailed face analysis, AR filters, high-precision applications

Future Work:

- Evaluate on more diverse datasets (different lighting, poses, occlusions)
- Compare with newer facial landmark models as they become available