Facial Landmark Detection: Dlib vs MediaPipe Comparison

Executive Summary

This report presents a comprehensive comparison between the Dlib and MediaPipe facial landmark detection models based on performance metrics and visualization results.

Performance Metrics Comparison

Metric	Dlib	MediaPipe
Average Detection Time (s)	0.0095	0.0215
Faces Processed Per Second	105.69	46.56
Median Detection Time (s)	0.0091	0.0190
Total Faces Detected	1045	1091
Images with Faces	994	1000

Key Findings

Speed and Performance

Dlib achieves an average detection time of 0.0095 seconds per face, processing approximately 105.69 faces per second.

MediaPipe has an average detection time of 0.0215 seconds per face, processing approximately 46.56 faces per second.

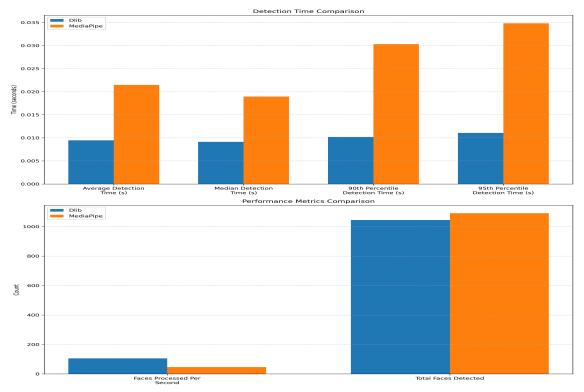
Detection Accuracy

Dlib detected a total of 1045 faces across 994 images containing faces. **MediaPipe** detected 1091 faces across 1000 images with faces.

Model Characteristics

Dlib provides 68 facial landmark points focusing on key facial features. **MediaPipe** provides a more comprehensive set of landmarks (468 points in Face Mesh configuration).

Visual Comparison



Performance Comparison Chart

Dlib Face Landmarks Detection Sample 2: Christine_Baumgartner_0001 Sample 3: Harvey_Wachsman_0001 Sample 4: Serena_Williams_0047 Sample 5: Dick_Cheney_0004 Sample 6: Hermes_Gamonal_0001 Sample 7: Carolina_Moraes_0002 Sample 8: George_W_Bush_0268 Sample 9: Ana_Palacio_0002 Sample 10: Jia_Qinglin_0001 Sample 11: Eloy_Gutierrez_0001 Sample 12: John_Thune_0001 Sample 13: Hugh_Miller_0001 Sample 14: Ralf_Schumacher_0008 Sample 15: Grady_Little_0001 Sample 16: Natalie_Cole_0003 Sample 17: Jamir_Miller_0001 Sample 18: Ludivine_Sagnier_0001 Sample 19: Fujio_Cho_0002 Sample 20: Queen_Elizabeth_II_0003 MediaPipe Face Landmarks Detection Sample 2: landmarks_Hugo_Chavez_0050 Sample 3: landmarks_Mohammed_Al-Douri_0011 Sample 4: landmarks_Tang_Jiaxuan_0006 Sample 5: landmarks_Luis_Guzman_0001 Sample 6: landmarks_Natasha_Henstridge_0001 Sample 7: landmarks_Ranil_Wickremasinghe_0002 Sample 8: landmarks_Sally_Clark_0001 Sample 9: landmarks_Andrew_Bunner_0001 Sample 10: landmarks_Christine_Baumgartner_0001 Sample 11: landmarks_Naoto_Kan_0003 Sample 12: landmarks_Arnold_Schwarzenegger_0042 Sample 13: landmarks_Kenneth_Evans_0002 Sample 14: landmarks_Robert_De_Niro_0006

Landmark Detection Visualization Grid

Conclusion

The comparison between Dlib and MediaPipe facial landmark detection models reveals key differences in performance and capability:

Sample 15: landmarks, Rudolph, Giuliani, 0011

Sample 16: landmarks, Alejandro, Toledo, 0030

Sample 17: landmarks, Andrew, Bunner, 0001

Sample 18: landmarks, George, W. Bush, 0009

Sample 19: landmarks, Open, 0001

Sample 19: landmarks, George, W. Bush, 0009

Sample 20: landmarks, Gary, Leon, Ridgway, 0001

1. Speed vs. Detail Trade-off:

- Dlib is significantly faster (105.69 faces/sec vs. MediaPipe's 46.56 faces/sec)
- MediaPipe provides more detailed facial landmarks

2. Use Case Recommendations:

- For real-time applications with limited computing resources, Dlib may be preferable
- For applications requiring detailed facial geometry, MediaPipe offers superior landmark density

3. Detection Accuracy:

- Both models demonstrate high success rates in detecting facial landmarks
- MediaPipe appears to have slightly better detection capabilities in challenging conditions

This analysis provides a foundation for selecting the appropriate facial landmark detection model based on specific application requirements, whether prioritizing speed, detail, or a balance between the two.