



## **Evaluasi Jarak Deteksi Antara Gestur Tangan Dan Kamera Webcam Dengan Metode Mediapipe**

**Julius Sembiring<sup>1,1</sup>, Vara Susilowati<sup>1,2</sup>, Vinsensius Reinard<sup>1,3</sup>, Meirista Wulandari<sup>1,4</sup>**  
<sup>1,2,3,4</sup>Program Studi Teknik Elektro, Fakultas Teknik, Universitas Tarumanagara, Jakarta Barat, DKI Jakarta, Indonesia, 11440  
Email: [julius.525210014@stu.untar.ac.id](mailto:julius.525210014@stu.untar.ac.id),

### **ABSTRAK**

Pengenalan gerakan dan interaksi manusia-mesin telah menjadi fokus utama dalam pengembangan teknologi. Aplikasi yang melibatkan gerakan tangan sebagai pengendali semakin berkembang, dari realitas virtual hingga perangkat lunak interaktif. Salah satu perkembangan penting dalam teknologi pendeteksian tangan adalah kerangka kerja MediaPipe dari Google. Menggunakan pendekatan pemrosesan citra dan pengenalan pola, MediaPipe mampu mengidentifikasi dan melacak gerakan tangan secara real-time, memiliki potensi aplikasi yang luas. Tantangan utama dalam menerapkan teknologi ini adalah evaluasi kinerja sistem dalam mengukur jarak antara tangan pengguna dan perangkat pendeteksian. Akurasi pengukuran jarak memengaruhi respons sistem terhadap gerakan tangan dan interaksi yang diinginkan. Variabilitas kondisi lingkungan, pencahayaan, dan posisi tangan dapat mempengaruhi kinerja teknologi pendeteksian jarak. Oleh karena itu, evaluasi MediaPipe terhadap deteksi jarak menjadi penting untuk memahami kemampuannya secara konsisten dan akurat. Hasil pengujian menunjukkan bahwa webcam dapat mendeteksi gestur tangan pada rentang jarak 20 cm – 290 cm. Namun, di atas 290 cm – 400 cm, beberapa jenis gestur tangan mulai tidak terdeteksi. Pada jarak di atas 450 cm, semua jenis gestur tangan tidak dapat terdeteksi oleh kamera webcam.

**Kata Kunci** : MediaPipe, Gestur tangan, Jarak, Deteksi.

### **ABSTRACT**

*In recent years, the emphasis on gesture recognition and human-machine interaction has surged, with hand gestures serving as controllers in applications ranging from virtual reality to interactive software. Central to this trend is Google's MediaPipe framework, a pivotal development in hand detection technology. Leveraging image processing and pattern recognition, MediaPipe achieves real-time identification and tracking of hand gestures, promising broad applications. However, an essential challenge lies in evaluating the system's proficiency in measuring the distance between the user's hand and the detection device, a factor crucial for a natural and intuitive user experience. The assessment of MediaPipe reveals that the webcam can adeptly detect hand gestures within a distance range of 20 cm to 290 cm. Beyond 290 cm up to 400 cm, certain hand gestures start eluding detection. Notably, at distances surpassing 450 cm, all types of hand gestures become undetectable by the webcam camera. Understanding these distance limitations is imperative, considering the impact of environmental conditions, lighting variations, and hand positioning on the accuracy of distance detection technology. This nuanced evaluation provides insights into MediaPipe's capabilities, enabling informed decisions for applications requiring consistent and precise hand gesture interactions.*

**Keywords:** MediaPipe, Hand gesture, Distance, Detection.

Submitted : ..... Revision : ..... Accepted : .....

## PENDAHULUAN

Pengenalan gerakan manusia sudah pernah diidentifikasi menggunakan CCTV, Robot, dan Interaksi Manusia Komputer (IMK). Seiring berjalannya kemajuan teknologi, pengenalan tersebut diterapkan menggunakan keilmuan bidang *Computer Vision* dan *Machine Learning*. Konektivitas teknologi digital yang semakin meningkat menjadikan peran teknologi semakin penting dalam pelaksanaan aktivitas sehari-hari. Teknologi mendukung berbagai aspek kehidupan kita, termasuk menyelesaikan tugas-tugas, melakukan pekerjaan, berbelanja, berkomunikasi, hiburan, serta mencari informasi atau berita. Aplikasi seperti permainan interaktif, realitas virtual, dan antarmuka pengguna berbasis gerak semakin populer, menggambarkan pentingnya teknologi pendeteksian tangan dan gerak tubuh dalam menciptakan pengalaman yang lebih imersif dan alami.

Gestur merupakan bentuk komunikasi non-verbal dengan aksi tubuh yang mengkomunikasikan pesan-pesan tertentu, sebagai pengganti ucapan. Sistem pengenalan gestur tangan terus berkembang, terutama untuk menggantikan peran perangkat masukan seperti mouse, keyboard dan *joystick* yang digunakan pada interaksi manusia dengan komputer, namun perangkat ini memiliki kelemahan diantaranya tidak alami bagi manusia dan jarak pengontrolan masih cukup dekat. Sistem pengenalan gestur tangan mengimplementasikan cara kerja sistem penglihatan dan pengenalan yang melibatkan mata manusia sebagai sensor yang merekam objek berupa gestur tangan dan kemudian mengirim informasi ke otak untuk diproses sehingga menghasilkan sebuah *output*.

Salah satu alat yang banyak digunakan untuk pendeteksian tangan adalah *Machine Learning* MediaPipe. MediaPipe dapat digunakan sebagai *pipeline* untuk melakukan

inferensi dari data citra. MediaPipe menggunakan model *machine learning* dengan akurasi tinggi untuk pendeteksian dan pelacakan, melacak titik titik pada bagian tubuh atau disebut juga *landmarks* [4]. Penggunaan MediaPipe dalam beberapa penelitian juga telah menunjukkan hasil yang menjanjikan. Sebagai contoh, penelitian oleh [5] menggunakan pustaka MediaPipe untuk Penerapan metode *hand gesture recognition* dalam melakukan kontrol terhadap aplikasi powerpoint dan media player untuk kebutuhan *online conference*. Penelitian lainnya, seperti yang dilakukan oleh Lugaresi *et al* tentang MediaPipe dalam kerangka kerja membangun persepsi objek [6].

Pengenalan *Hand Gesture* dapat dilakukan dengan mendeteksi dan melacak posisi jari dan telapak tangan. Pengenalan gerakan tangan dibagi menjadi 2 jenis, yaitu statis dengan membentuk gerakan tangan saja, dan dinamis dengan membentuk isyarat tangan dan kemudian menggerakannya [7]. Tujuan utama dalam mempelajari gestur tangan adalah untuk memperkenalkan sistem yang dapat mendeteksi gerak tubuh manusia dan menggunakannya untuk menyampaikan informasi atau untuk tujuan perintah dan kontrol [8].

Beberapa jenis pendekatan didasarkan pada teknik visi komputer konvensional, seperti pencocokan pola dan ekstraksi fitur. Metode ini mengandalkan fitur manual dan sering kali memerlukan banyak pengetahuan domain serta penyetulan manual untuk mendapatkan performa yang baik. Pendekatan lain mengandalkan teknik pembelajaran mendalam, seperti jaringan saraf tiruan dan jaringan saraf berulang. Metode ini telah mencapai kinerja tingkat lanjut dalam tugas deteksi tangan dan pengenalan gerakan. Contoh dari sistem tersebut adalah *MediaPipe Hands*, yang menggunakan data *RGB* untuk mengidentifikasi *landmark* dari tangan dan melacak pergerakannya [9].

MediaPipe *hand* merupakan solusi dari pelacakan atau pendeteksian tangan dan jari yang memiliki ketelitian tinggi. MediaPipe memanfaatkan pembelajaran *Machine Learning* (ML) untuk menyimpulkan 21 3D *landmark* dan diambil dari 1 *frame*. Contoh *Landmark* yang dapat diambil dari 1 *frame* ini diperlihatkan pada Gambar 1. MediaPipe melakukan pelatihan detektor tangan [5].



Gambar 1. Hand Landmark MediaPipe [5]

MediaPipe ditujukan untuk pelacakan gerakan yang dilakukan oleh seseorang. Pada setiap anggota tubuh ditampilkan titik-titik yang dihubungkan oleh garis. Pada garis tersebut akan menghubungkan titik-titik koordinat dan membentuk kerangka seseorang. *Library* ini harus dieksekusi sekalian dengan *library OpenCV* agar video dari gambar pelacakan dapat ditampilkan [10]. MediaPipe telah menunjukkan potensi yang menjanjikan dalam deteksi dan pelacakan anggota tubuh manusia. Maka dari itu penelitian ini akan memfokuskan pada evaluasi MediaPipe terhadap jarak dengan memanfaatkan anggota tubuh yaitu tangan agar dapat terdeteksi. Tujuan dari penelitian ini adalah untuk mengevaluasi jarak tangan yang mampu terdeteksi oleh MediaPipe sehingga didapatkan jangkauan jarak optimal.

## METODOLOGI

MediaPipe sebagai *Framework* untuk membangun *pipeline* pembelajaran mesin untuk memproses data deret waktu seperti video, audio, dll. *Framework* lintas platform ini bekerja di desktop/server, Android, iOS, dan perangkat tertanam seperti *Raspberry PI* dan lain-lain [11]. MediaPipe terdiri dari tiga bagian utama: (a) kerangka kerja untuk menyimpulkan dari data sensorik, (b) seperangkat alat untuk evaluasi kinerja, dan (c) kumpulan komponen pemrosesan dan

kesimpulan yang dapat digunakan kembali yang disebut kalkulator.

MediaPipe berhubungan dengan OpenCV dikarenakan dari salah satu fitur dari OpenCV yakni seperti deteksi wajah dan deteksi objek [12], MediaPipe dan OpenCV dapat digunakan bersama untuk menangani data video. Bahasa pemrograman yang digunakan dalam analisa ini adalah bahasa pemrograman python, dimana python dapat digunakan untuk menulis skrip yang menggunakan MediaPipe dan OpenCV untuk melakukan berbagai tugas visi komputer, seperti deteksi objek seperti gestur tangan.

Dalam memulai pemrograman untuk deteksi gestur tangan berbasis mediapipe ini, proses instalasi *library* dan *environment* diperlukan. Pada proses instalasi *library* yang digunakan yakni MediaPipe dan OpenCV. Proses instalasi MediaPipe adalah sebagai berikut; gunakan pip. Jalankan ke command prompt "pip install mediapipe" untuk menginstal versi terbaru dari MediaPipe sedangkan untuk penginstalan OpenCV sebagai berikut; gunakan pip, jalankan ke command prompt "pip install opencv-python". Pada proses instalasi *environment* yang digunakan yakni bahasa pemrograman python. Proses instalasi python dapat mengikuti tahapannya yang ada di alamat web, <https://www.python.org/downloads>.

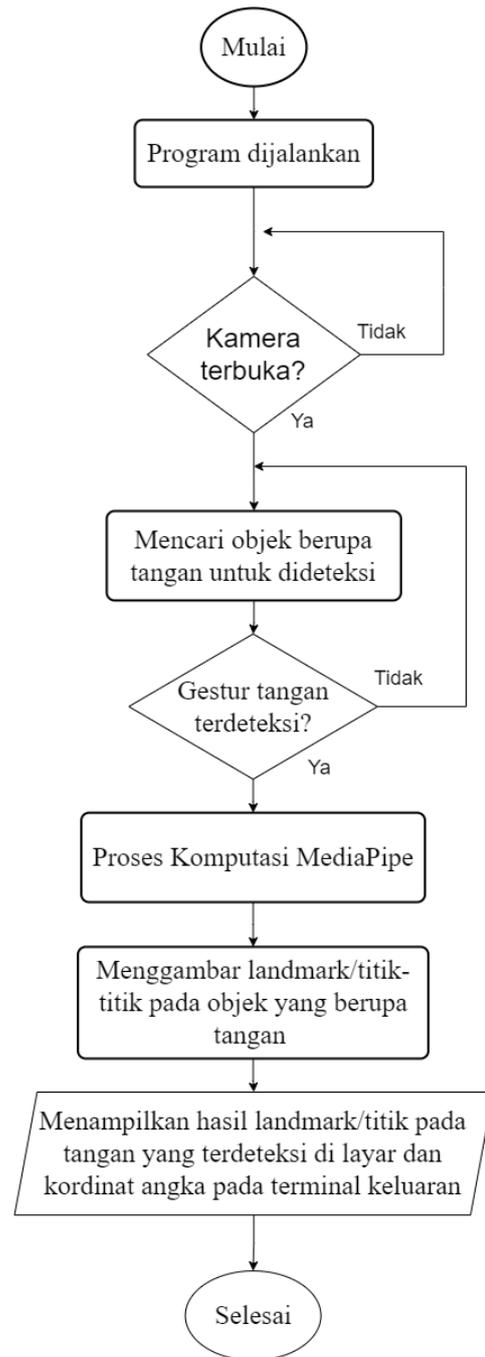
Beberapa jenis pendekatan didasarkan pada teknik visi komputer konvensional, seperti pencocokan pola dan ekstraksi fitur. Metode ini mengandalkan fitur manual dan sering kali memerlukan banyak pengetahuan domain serta pengaturan manual untuk mendapatkan performa yang baik. Pendekatan lain mengandalkan teknik pembelajaran mendalam, seperti jaringan saraf tiruan dan jaringan saraf berulang. Metode ini telah mencapai kinerja tingkat lanjut dalam tugas deteksi tangan dan pengenalan gerakan. Contoh dari sistem tersebut adalah *MediaPipe Hands*, yang menggunakan data RGB untuk mengidentifikasi *landmark* dari tangan dan melacak pergerakannya [9].

Metode penelitian ini yaitu menggunakan pipeline MediaPipe untuk melakukan pengenalan gerakan tangan secara statis, serta mengetahui jarak jangkauan *gesture* tangan yang dapat terdeteksi oleh MediaPipe. Proses deteksi tangan dari MediaPipe dapat dilihat pada Gambar 2. Hal ini dimaksudkan untuk mengetahui performa terbaik dari sistem MediaPipe. Performa yang dimaksud adalah kemampuan MediaPipe melakukan estimasi akurasi terkait jarak antara kamera webcam dan pengguna.



Gambar 2. Alur Sistem Deteksi Gestur Tangan

Proses dari pengenalan gestur tangan dimulai dari membuka kamera streaming yakni kamera webcam. Setelah kamera webcam terbuka, sistem menerima *input* berupa data gambar tangan yang terdeteksi oleh kamera webcam. Proses pendeteksian tangan dilakukan dengan perintah `VideoCapture`. Perintah `VideoCapture` ini digunakan untuk mencari fitur-fitur tangan dengan bantuan perintah proses pengolahan data warna. Ketika data warna sudah sesuai, data tangan tersebut diolah untuk menggambar *hand landmark* MediaPipe. *Hand landmark* digambarkan dengan garis berwarna hijau beserta dengan titik-titik merah dalam satu frame. Hasil pembentukan *hand landmark* itu menandakan bahwa kamera berhasil mendeteksi objek gestur tangan. Koordinat setiap titik pada tangan disimpan oleh program. Adapun tampilan diagram alir pada setiap proses pemrograman MediaPipe yang dapat dilihat pada Gambar 3.



Gambar 3. Diagram Alir Deteksi Gestur Tangan dengan Metode MediaPipe

## HASIL DAN PEMBAHASAN

Berbagai jenis gestur tangan dan jarak diuji dalam penelitian ini. Gestur tangan yang diuji memperlihatkan beberapa pola seperti satu jari, dua jari, tiga jari, empat jari, lima jari dan dua tangan seperti terlihat pada Gambar 4. Data gestur tangan tersebut menjadi *input* dalam

evaluasi jarak MediaPipe. Pada evaluasi ini membahas tentang jangkauan jarak terhadap gestur tangan yang dapat terdeteksi atau tidak terdeteksi oleh MediaPipe.



Gambar 4. Data Gestur Tangan yang Diuji

Pengujian dibagi menjadi dua yakni untuk pengujian pertama membahas tentang jangkauan jarak pengujian terhadap data. range yang telah ditentukan seperti pada Tabel 1. Jarak telapak tangan terhadap kamera diukur dari 5 cm hingga 450 cm. Pada umumnya, gestur tangan mulai dapat terdeteksi pada jarak 20 cm. Namun jangkauan jarak terjauh yang dapat terdeteksi oleh MediaPipe terhadap gestur tangan berbeda antara satu gestur dengan gestur lainnya. Pada gestur tangan 1 jari, gestur tangan 1 jari tidak dapat terdeteksi saat jarak 350 cm. Pada gestur tangan 2 jari dan gestur tangan 3 jari, gestur tangan mulai tidak terdeteksi pada jarak 400 cm. Pada gestur tangan 4 jari, gestur tangan 5 jari, dan gestur dua tangan, gestur tangan mulai tidak terdeteksi pada jarak 450 cm.

Tabel 1. Uji Keberhasilan MediaPipe terhadap Jarak Berskala dari Kamera Webcam

Gestur Tangan	Jarak (cm)											
	5	10	20	40	70	110	160	220	290	350	400	450
1 Jari	x	x	v	v	v	v	v	v	v	x	x	x
2 Jari	x	x	v	v	v	v	v	v	v	v	x	x
3 Jari	x	x	v	v	v	v	v	v	v	v	x	x
4 Jari	x	x	v	v	v	v	v	v	v	v	v	x
5 Jari	x	x	v	v	v	v	v	v	v	v	v	x
2 Tangan	x	x	v	v	v	v	v	v	v	v	v	x

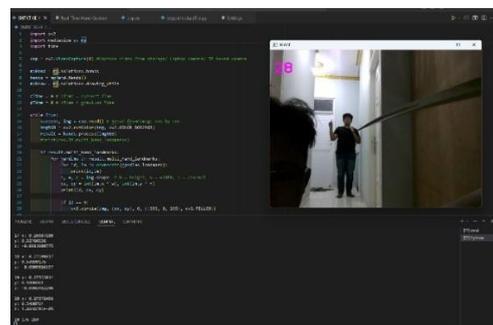
\*x = tidak berhasil  
v = berhasil

Pengujian kedua dilakukan untuk menguji rata-rata jarak minimum dan rata-rata jarak maksimum kemampuan MediaPipe mendeteksi kelima gestur tangan. Terdapat 3 kali percobaan yang dilakukan. Percobaan dilakukan masing-masing pada jarak minimum dan jarak maksimum kemudian nilai rata-rata jarak dihitung. Hasil pengujian kedua dapat dilihat pada Tabel 2. Tabel 2 menunjukkan bahwa rata-rata jarak minimum pada gestur tangan dan webcam berkisar antara 21 cm sampai dengan 26,3 cm. Sedangkan, rata-rata jarak maksimum pada gestur tangan dan webcam berkisar antara 342,7 cm sampai dengan 442,7 cm.

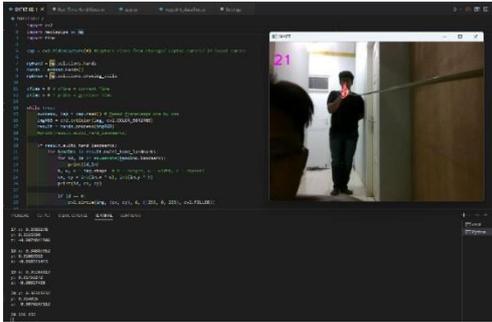
Tabel 2. Uji Deteksi MediaPipe pada Jarak Maksimal dan Jarak Minimal

Gestur Tangan	Rata-rata jarak minimum (cm)	Rata-rata jarak maksimum (cm)
1 Jari	21	342,7
2 Jari	22,6	358,3
3 Jari	23,3	383
4 Jari	25	412,7
5 Jari	24,7	434
2 Tangan	26,3	442,7

Saat salah satu gestur tangan berada melebihi jarak maksimal nya maka tidak akan terdeteksi seperti pada Gambar 5(a) dan sebaliknya saat salah satu gestur tangan masih berada di jarak maksimal nya maka akan tetap terdeteksi seperti pada Gambar 5(b).



Gambar 5(a). Contoh Tampilan Gestur Tangan yang Berhasil



Gambar 5(b). Contoh Tampilan Gestur Tangan yang Tidak Berhasil

- I. Suyudi *et al.*, “Pengenalan Bahasa Isyarat Indonesia menggunakan Mediapipe dengan Model Random Forest dan Multinomial Logistic Regression (Introduction to Indonesian Sign Language Using Mediapipe With Random Logistic Forest Models and Multinomial Logistic Regression),” *Jurnal Ilmu Siber dan*

## KESIMPULAN

Penelitian evaluasi semacam ini akan memberikan wawasan mendalam tentang kemampuan MediaPipe dalam mengukur jarak secara konsisten dan akurat. Dengan menguji sistem dalam berbagai situasi dan skenario, evaluasi ini akan membantu mengidentifikasi kelebihan dan keterbatasan teknologi, serta memberikan panduan bagi pengembangan lebih lanjut. Selain itu, jarak ideal untuk semua gestur tangan dapat terdeteksi yakni pada saat jarak bernilai 20 cm sampai dengan 290 cm dari kamera webcam. Jarak yang diukur hasil pengujian ini membuktikan bahwa dengan jarak 20 cm – 290 cm, semua jenis gestur tangan yang di uji dapat terdeteksi. Untuk jarak di atas 290 cm – 400 cm, mulai terlihat beberapa jenis gestur tangan yang mulai tidak terdeteksi. Untuk jarak di atas 450 cm, semua jenis gestur tangan tidak dapat terdeteksi oleh kamera webcam.

## DAFTAR PUSTAKA

- F. Daniel Tanugraha, H. Pratikno, M. Musayanah, and W. Indah Kusumawati, “Pengenalan Gerakan Olahraga Berbasis (Long Short-Term Memory) Menggunakan Mediapipe,” *Journal of Advances in Information and Industrial Technology*, vol. 4, no. 1, pp. 37–45, May 2022, doi: 10.52435/jaiit.v4i1.182.
- A. Suryaperdana Agoes, “Applying Hand Gesture Recognition for User Guide Application Using MediaPipe,” 2021.
- S. Makahaube, A. M. Sambul, and S. R. U. A. Sompie, “Implementation of Gesture Recognition Technology for Self-Education Service Platform,” *Jurnal Teknik Informatika*, vol. 16, no. 4, pp. 465–472, 2021.

- Teknologi Digital (JISTED)*, vol. 1, no. 1, pp. 65–80, 2022, doi: 10.35912/jisted.v1i1.1899.
- W. S. Wiyogo and L. Liliana, “Penerapan metode hand gesture recognition dalam melakukan kontrol terhadap aplikasi powerpoint dan media player untuk kebutuhan online conference,” *Jurnal Infra*, vol. 10, no. 2, pp. 167–171, 2022.
- C. Lugaresi *et al.*, “MediaPipe: A Framework for Building Perception Pipelines,” Jun. 2019, [Online]. Available: <http://arxiv.org/abs/1906.08172>
- H. Marcellino, A. Yudo Husodo, and F. Bimantoro, “DEVELOPMENT OF HAND GESTURE COMMAND USING LEAP MOTION FOR FLOORPLANNER APP.”
- M. Oudah, A. Al-Naji, and J. Chahl, “Hand Gesture Recognition Based on Computer Vision: A Review of Techniques,” *Journal of Imaging*, vol. 6, no. 8. MDPI, Jul. 01, 2020. doi: 10.3390/JIMAGING6080073.
- A. Vysocký, T. Poštulka, J. Chlebek, T. Kot, J. Maslowski, and S. Grushko, “Hand Gesture Interface for Robot Path Definition in Collaborative Applications: Implementation and Comparative Study,” *Sensors*, vol. 23, no. 9, May 2023, doi: 10.3390/s23094219.
- M. Abdul muthalib, I. Irfan, K. Kartika, and S. M. Selamat Meliala, “PENGIRAAN POSE MODEL MANUSIA PADA REPETISI KEBUGARAN AI PEMOGRAMAN PYTHON BERBASIS KOMPUTERISASI,” *INFOTECH journal*, vol. 9, no. 1, pp. 11–19, Jan. 2023, doi: 10.31949/infotech.v9i1.4233.
- R. Kulhary, “OpenCV Overview,” *GeeksforGeeks*, Sep. 23, 2019. Accessed: Sep. 06, 2023. [Online]. Available: <https://www.geeksforgeeks.org/opencv-overview/>
- Kukil, “Introduction to MediaPipe,” *LearnOpenCV – Learn OpenCV, PyTorch, Keras, Tensorflow with examples and tutorials*, Mar. 01, 2022. <https://learnopencv.com/introduction-to-mediapipe/#What-is-MediaPipe?>.