



Pengembangan Sistem Kendali Lengan Robot Humanoid Berbasis Pengolahan Citra Real-Time Menggunakan MediaPipe

Vemas Nandra Lupito^{1a}, Muchamad Malik¹, Aan Burhanudin¹

¹Program Studi Teknik Mesin, Fakultas Teknik dan Informatika, Universitas PGRI Semarang

Korespondensi:

^aProgram Studi Teknik Mesin, Fakultas Teknik dan informatika, Universitas PGRI Semarang
vemasnandra@gmail.com

ABSTRAK

Kemajuan teknologi dalam bidang pengolahan citra telah mendorong inovasi pada sistem kendali robot humanoid yang lebih responsif dan efisien. Penelitian ini bertujuan untuk mengembangkan sistem kendali lengan robot humanoid berbasis pengolahan citra *real-time* dengan memanfaatkan *framework* MediaPipe. Sistem dirancang untuk meniru gerakan lengan manusia secara langsung dengan menangkap data landmark tubuh melalui kamera laptop sebagai sensor visual utama. Koordinat titik-titik tubuh seperti bahu, siku, dan pergelangan tangan diproses menggunakan prinsip trigonometri berbasis vektor untuk menghitung sudut artikulasi, yang kemudian dikonversi menjadi sinyal kendali untuk motor servo. Struktur mekanik robot dibangun menggunakan bahan PLA hasil pencetakan 3D, dengan konfigurasi tiga derajat kebebasan pada bahu, siku, dan pergelangan tangan. Mikrokontroler ESP32 digunakan sebagai unit kendali utama melalui komunikasi serial dari laptop. Hasil pengujian menunjukkan bahwa sistem mampu meniru gerakan manusia dengan tingkat responsivitas yang baik. Gerakan *fleksi* pada siku mencapai 135°, setara dengan 93% dari rentang gerak manusia, sementara torsi maksimum yang dibebankan pada sambungan bahu sebesar 20 kg·cm masih berada dalam batas aman spesifikasi servo HS-805BB. Sistem menunjukkan kinerja optimal dalam lingkungan terkendali dengan latensi sekitar 120 ms. Dengan demikian, sistem ini berhasil merepresentasikan prototipe awal *mirror-based teleoperation* berbasis visi komputer yang efisien dan ekonomis. Penelitian ini berpotensi dikembangkan lebih lanjut untuk aplikasi di bidang robot asistif, rehabilitasi, dan interaksi manusia-robot.

Kata kunci: MediaPipe, lengan robot humanoid, pengolahan citra, servo, ESP32, Python.

ABSTRACT

Advancements in image processing technology have driven innovation in humanoid robot control systems that are more responsive and efficient. This study aims to develop a humanoid robot arm control system based on real-time image processing using the MediaPipe framework. The system is designed to imitate human arm movements by capturing body landmark data through a laptop camera as the primary visual sensor. The coordinates of body points such as the shoulder, elbow, and wrist are processed using vector-based trigonometric calculations to determine joint angles, which are then converted into servo motor control signals. The robot's mechanical structure is constructed using 3D-printed PLA material, with three degrees of freedom corresponding to the shoulder, elbow, and wrist joints. An ESP32 microcontroller is employed as the main control unit via serial communication from a laptop. Experimental results demonstrate that the system is capable of accurately replicating human movements with good responsiveness. Elbow flexion reached up to 135°, equivalent to 93% of the natural human range, while the maximum torque load on the shoulder joint was

20 kg-cm, remaining within the safe operating limits of the HS-805BB servo motor. The system performs optimally in a controlled environment with an average latency of approximately 120 ms. In conclusion, the system successfully represents an early-stage prototype of a vision-based mirror-based teleoperation platform that is both efficient and cost-effective. This research holds promise for further development in assistive robotics, rehabilitation systems, and human-robot interaction based on visual input.

Keywords: MediaPipe, humanoid robot arm, image processing, servo motor, ESP32, Python.

1. PENDAHULUAN

Perkembangan teknologi di bidang robotika menunjukkan kemajuan yang signifikan, terutama dalam inovasi robot humanoid yang dirancang untuk mampu berinteraksi secara langsung dengan manusia. Salah satu komponen penting dari robot humanoid adalah lengan robot, yang berfungsi sebagai alat interaksi fisik antara robot dan lingkungan sekitarnya. Kontrol lengan robot secara presisi dan *real-time* sangat diperlukan dalam berbagai aplikasi seperti layanan kesehatan, pendidikan, serta interaksi manusia-robot di bidang industri dan sosial. Teknologi *computer vision* menjadi solusi potensial dalam mengembangkan sistem kendali robot yang lebih alami dan intuitif, tanpa ketergantungan pada perangkat keras tambahan seperti sensor gerak atau pengendali manual. Salah satu pendekatan yang berkembang pesat adalah penggunaan MediaPipe, sebuah pustaka *machine learning* dari Google yang dirancang untuk memproses dan mendeteksi *pose* tubuh secara *real-time* melalui citra kamera. Dengan memanfaatkan MediaPipe, gerakan manusia dapat dikenali dan diinterpretasikan untuk mengontrol lengan robot secara langsung, sehingga meningkatkan efisiensi dan interaktivitas sistem. Dalam proses verifikasi pengendalian berjalan pada robot humanoid, efektivitas yang terbukti melalui simulasi tidak selalu dapat direalisasikan secara langsung pada implementasi di dunia nyata [1]. Ketidakesesuaian ini umumnya disebabkan oleh ketidakakuratan model, seperti *defleksi* struktural, keterbatasan kekakuan servo, serta kesalahan pengukuran sensor. Sistem robot memiliki tiga komponen dasar, yaitu: manipulator, kontroler, dan power (daya) [2,3]. MediaPipe merupakan sebuah kerangka kerja bersifat open-source yang dikembangkan oleh Google, dirancang untuk mendukung proses deteksi serta pelacakan tangan secara waktu nyata [4]. Dalam sektor industri, pemanfaatan robot telah memberikan kontribusi signifikan dalam mendukung efisiensi proses kerja di lingkungan pabrik [5,6].

Meskipun telah banyak penelitian yang membahas pengendalian robot humanoid, sebagian besar masih bergantung pada perangkat input tambahan seperti sensor inertial (IMU), joystick, atau perintah berbasis suara. Pendekatan-pendekatan tersebut memiliki keterbatasan dalam hal fleksibilitas, *real-time response*, dan implementasi pada lingkungan dinamis. Selain itu, penelitian terdahulu yang menggunakan pengolahan citra cenderung berfokus pada deteksi wajah atau isyarat tangan saja, tanpa integrasi penuh dengan sistem aktuasi robot secara langsung. Hingga saat ini, masih terbatas penelitian yang secara spesifik mengimplementasikan MediaPipe sebagai metode utama dalam mengendalikan gerakan lengan robot humanoid secara *real-time* dan presisi tinggi. Oleh karena itu, dibutuhkan penelitian yang mengisi celah ini dengan mengembangkan sistem kendali lengan robot yang terintegrasi dengan deteksi gerakan tubuh manusia berbasis *image processing*. Robot umumnya dimanfaatkan untuk menjalankan tugas-tugas yang bersifat berat, berbahaya, maupun repetitif, karena memiliki keunggulan dalam hal kecepatan dan akurasi yang tinggi [7,8]. Salah satu jenis robot yang banyak diterapkan dalam lini produksi adalah lengan robotik, yang berfungsi sebagai alat bantu pemindahan produk [9–11]. Dari perspektif ekonomi, penggunaan robot di industri memberikan sejumlah manfaat, antara lain penghematan biaya tenaga kerja serta peningkatan konsistensi dan kualitas hasil produksi [12–15]. Estimasi *pose* manusia dari data video memegang peranan krusial dalam berbagai aplikasi, di antaranya evaluasi aktivitas fisik, pengenalan bahasa isyarat, serta pengendalian gerakan tubuh secara keseluruhan [16–18].

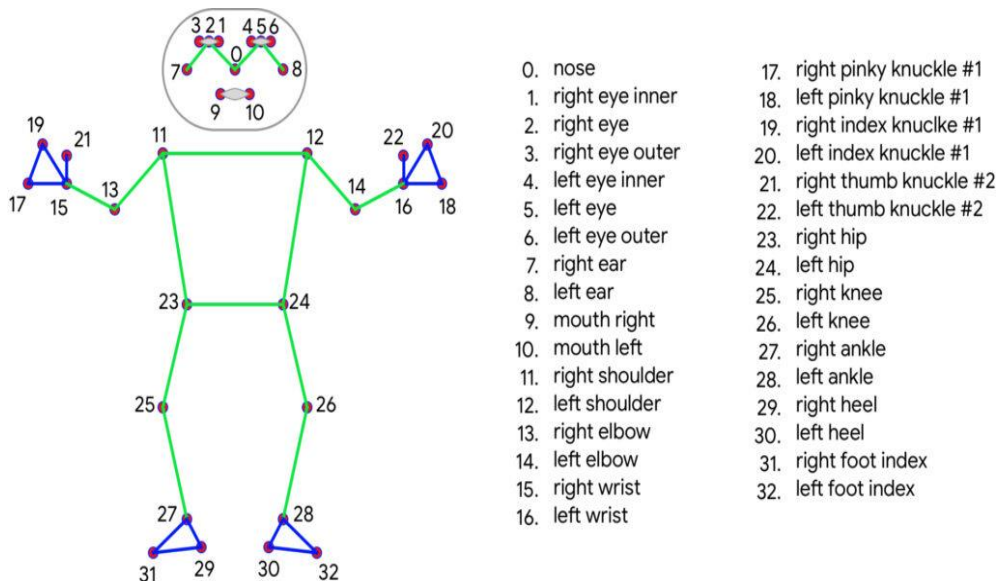
Penelitian ini bertujuan untuk mengembangkan sistem kendali lengan robot humanoid berbasis pengolahan citra *real-time* menggunakan MediaPipe. Secara khusus, tujuan dari penelitian ini adalah: Menghubungkan hasil deteksi gerakan dengan sistem aktuator lengan robot untuk mereplikasi gerakan manusia, menguji akurasi dan responsivitas sistem dalam mengeksekusi perintah gerakan secara tepat waktu, memberikan solusi alternatif yang efisien dan rendah biaya dalam pengendalian robot humanoid berbasis visi komputer.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan eksperimental dengan merancang dan mengimplementasikan sistem lengan robot humanoid yang dikendalikan melalui gerakan tangan manusia secara *real-time*. Sistem dikembangkan menggunakan *framework* MediaPipe yang dijalankan dengan bahasa pemrograman Python, memanfaatkan kamera internal laptop sebagai sensor input utama. Proses diawali dengan akuisisi citra tangan pengguna melalui kamera, yang kemudian diproses oleh modul *Hand Tracking* dari MediaPipe untuk mendeteksi dan melacak 21 titik *landmark* tangan.

Data koordinat dari *landmark* tersebut dikonversi menjadi sinyal kendali yang digunakan untuk menggerakkan aktuator pada lengan robot sesuai dengan posisi dan gerakan tangan pengguna. Validasi sistem dilakukan dengan mengamati tingkat akurasi dan responsivitas gerakan lengan robot terhadap variasi posisi tangan dalam kondisi pencahayaan yang berbeda. Seluruh eksperimen dilakukan dalam lingkungan terkontrol untuk memastikan konsistensi hasil pengujian.

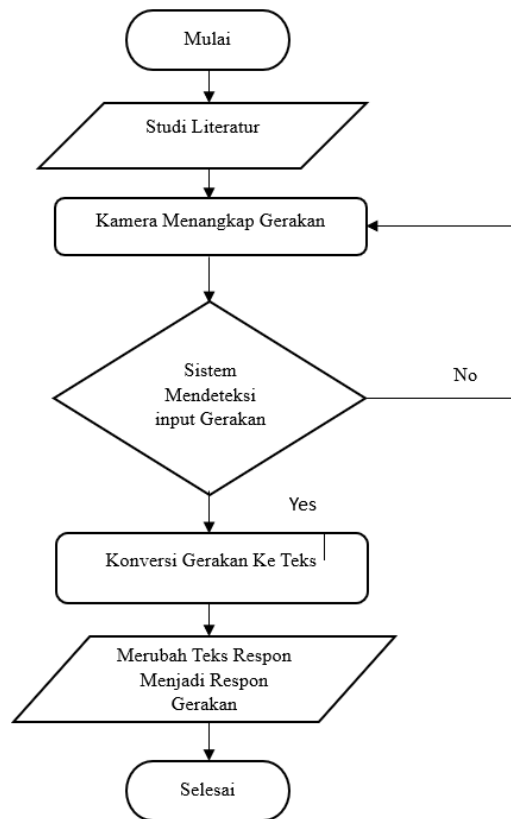
Deteksi tangan merupakan proses identifikasi keberadaan tangan dalam citra statis maupun video [19], sedangkan pelacakan tangan mengacu pada proses pemantauan pergerakan dan posisi tangan secara berkelanjutan dari satu *frame* ke *frame* berikutnya. Teknologi ini memiliki peran signifikan dalam berbagai aplikasi, khususnya dalam sistem pengenalan *gestur* serta interaksi antara manusia dan komputer. Salah satu solusi yang banyak digunakan adalah Mediapipe, yang menawarkan model deteksi tangan berbasis pembelajaran mesin yang telah dilatih menggunakan kumpulan data citra tangan yang komprehensif. Model ini mampu mendeteksi serta melacak pergerakan tangan dengan tingkat akurasi yang tinggi. Berbagai studi terdahulu menunjukkan bahwa teknologi deteksi dan pelacakan tangan memiliki potensi implementasi yang luas, antara lain pada sistem kendali berbasis isyarat, lingkungan realitas virtual (VR), serta aplikasi realitas tertambah (AR).



Gambar 1. Pose landmark

Gambar 1. Merupakan representasi skematik dari *pose landmark* model yang digunakan dalam sistem pelacakan dan pendeteksian postur tubuh manusia secara visual. Model ini terdiri atas 33 titik kunci (*keypoints*) yang mewakili bagian-bagian tubuh utama, termasuk kepala, tangan, badan, dan kaki. Setiap titik diberi nomor dan label yang menunjukkan letak anatomisnya, seperti hidung (0), mata (1–6), telinga (7–8), mulut (9–10), bahu (11–12), siku (13–14), pergelangan tangan (15–16), buku-buku jari (17–22), panggul (23–24), lutut (25–26), pergelangan kaki (27–28), tumit (29–30), serta ujung kaki (31–32). Titik-titik ini dihubungkan dengan garis berwarna (hijau dan biru) untuk merepresentasikan struktur rangka tubuh manusia secara visual. Model ini umumnya digunakan dalam sistem berbasis *computer vision* seperti MediaPipe, yang bertujuan untuk mengidentifikasi *pose* tubuh secara *real-time* pada gambar atau video. Informasi ini krusial dalam berbagai aplikasi seperti interaksi manusia-komputer, analisis olahraga, rehabilitasi medis, dan sistem kontrol robotik.

berbasis gerakan manusia. Desain visual pada gambar 1 menunjukkan keterhubungan antar titik kunci secara logis dan sistematis, sehingga memudahkan pengembangan algoritma pelacakan postur serta deteksi gerakan. Struktur model ini mencerminkan pendekatan berbasis *landmark* dalam memahami dinamika dan orientasi tubuh manusia.



Gambar 2. Diagram Alur Perancangan Software

Gambar 2. Merupakan diagram alur proses sistem deteksi dan konversi gerakan menjadi teks, yang banyak digunakan dalam penelitian interaksi manusia-komputer berbasis visi komputer. Diagram ini menggambarkan tahapan utama dalam sistem pengenalan gerakan yang bertujuan untuk menerjemahkan input visual menjadi keluaran berbasis teks. Tahapan pertama adalah Studi Literatur, yang menjadi landasan teoritis dalam mengembangkan sistem. Setelah itu, kamera digunakan untuk menangkap gerakan pengguna secara real-time. Gerakan yang tertangkap kemudian diproses oleh sistem untuk mendeteksi apakah terdapat input gerakan. Jika tidak terdeteksi, kamera akan terus menangkap gerakan hingga sistem menerima input yang valid. Ketika gerakan berhasil terdeteksi, sistem akan melakukan konversi gerakan ke dalam bentuk teks, yang merepresentasikan makna atau instruksi dari gerakan tersebut. Selanjutnya, teks hasil konversi diubah menjadi respon gerakan yang dapat digunakan untuk memberikan umpan balik terhadap pengguna atau sistem lain.

2.1 Batasan-Batasan Penelitian

Dalam rangka menjaga fokus dan efektivitas pelaksanaan penelitian, serta memastikan keterukuran dan keterlaksanaan dari pengembangan sistem, penelitian ini dibatasi oleh beberapa ruang lingkup berikut:

1. Deteksi Gerakan Terbatas pada Lengan dan Tangan

Penelitian ini hanya memfokuskan deteksi gerakan tubuh pada bagian lengan dan tangan. Komponen tubuh lain seperti kaki, kepala, dan torso tidak dianalisis secara menyeluruh. Fokus ini ditetapkan untuk menyesuaikan dengan tujuan utama, yaitu pengendalian lengan robot humanoid.

2. Lingkup Deteksi Pose Menggunakan MediaPipe

Sistem hanya menggunakan *MediaPipe Pose* dan/atau *MediaPipe Holistic* sebagai metode utama untuk deteksi pose. Tidak digunakan metode pendeteksian lain seperti OpenPose, Kinect, atau sensor IMU agar dapat mengevaluasi efektivitas MediaPipe secara terfokus.

3. **Lingkungan Uji yang Terkontrol**

Sistem diuji dalam lingkungan yang terkendali (indoor) dengan pencahayaan cukup dan latar belakang yang minim gangguan visual. Kondisi ini dipilih untuk memastikan stabilitas deteksi citra, mengingat sistem berbasis pengolahan citra sangat sensitif terhadap pencahayaan dan objek latar.

4. **Gerakan Sederhana dan Tidak Kompleks**

Gerakan yang dikenali dan diterjemahkan oleh sistem terbatas pada gerakan sederhana (misalnya: mengangkat tangan, menggerakkan tangan ke kiri atau kanan). Penelitian ini tidak mencakup pengenalan gerakan kompleks atau gerakan berurutan seperti bahasa isyarat penuh.

5. **Aktuator Robot Terbatas pada Dua DOF (*Degree of Freedom*)**

Lengan robot yang digunakan dalam penelitian ini memiliki derajat kebebasan terbatas (misalnya 2 DOF: bahu dan siku). Implementasi tidak mencakup aktuator dengan kompleksitas tinggi seperti jari atau pergelangan tangan multi-sumbu.

6. **Platform Kamera Menggunakan Kamera Laptop atau Webcam Eksternal**

Sistem hanya memanfaatkan kamera laptop atau webcam eksternal standar untuk akuisisi data visual. Penggunaan kamera industri atau kamera depth sensor tidak termasuk dalam cakupan penelitian ini.

7. **Tidak Menganalisis Aspek Kognitif atau Respons Emosional Robot**

Penelitian ini hanya berfokus pada aspek pengendalian fisik lengan robot, dan belum mencakup aspek interaksi kognitif atau respons emosional dari sistem robot terhadap pengguna.

3. HASIL DAN PEMBAHASAN

Implementasi mekanik pada sistem robot humanoid dalam penelitian ini dikombinasikan dengan pendekatan *vision-based control* menggunakan *framework* MediaPipe berbasis Python. Prototipe robot yang terdiri dari bagian torso dan dua lengan mekanik dirancang untuk meniru gerakan manusia secara *real-time* dengan masukan *visual* dari kamera laptop. Kamera laptop bertindak sebagai sistem penglihatan utama yang menangkap *pose* tubuh manusia, kemudian diterjemahkan ke dalam sinyal kontrol servo motor melalui pemrosesan citra menggunakan model *MediaPipe Pose Tracking* dari Google.



Gambar 3. Diagram Alur Deteksi Objek secara Real Time

Diagram alir tersebut menggambarkan suatu sistem infrastruktur terstruktur dan sistematis untuk deteksi objek secara *real-time* dan kontrol robot menggunakan penglihatan komputer dan komunikasi serial dengan mikrokontroler Arduino. Sistem dimulai dengan fase inisialisasi, di mana koneksi serial ke Arduino dibangun pada baud rate 9600, dan kamera *default* (indeks 0) diaktifkan. Fase ini memastikan semua komponen perangkat keras terkonfigurasi dengan benar sebelum masuk ke *loop* kontrol utama, yang sangat penting untuk operasi sistem yang andal [20]. Kamera menangkap citra secara kontinu dan setiap *frame* yang diperoleh dikonversi ke ruang warna HSV. Ruang warna ini dipilih karena lebih tahan terhadap variasi pencahayaan, menjadikannya lebih andal dalam segmentasi warna. Sistem kemudian menerapkan *masking* pada rentang HSV tertentu ([30,150,50] – [85,255,255]) untuk menyoroti objek dengan warna spesifik, dan menggunakan algoritma pencarian kontur untuk mendeteksi keberadaan objek.

Fungsi inti sistem berpusat pada *loop* pemrosesan video yang berkesinambungan. Setiap *frame* yang diambil dari kamera dikonversi ke ruang warna HSV, kemudian diterapkan filter rentang warna untuk mengisolasi objek target. Deteksi kontur digunakan untuk mengidentifikasi batas objek, dengan kontur terbesar dipilih untuk menentukan posisi objek. Persegi panjang pembatas digambar di sekitar objek yang terdeteksi sebagai umpan balik visual, suatu praktik umum dalam sistem penglihatan komputer untuk memverifikasi akurasi deteksi [15].

Logika kontrol robot diaktifkan berdasarkan posisi relatif objek terhadap pusat *frame*. Jika objek berada di sebelah kiri pusat dengan ambang batas 50 piksel, perintah ('L') dikirim ke Arduino untuk menggerakkan robot

ke kiri. Sebaliknya, jika objek berada di sebelah kanan, perintah ('R') dikirim untuk bergerak ke kanan. Ketika objek berada di tengah, perintah berhenti ('S') diberikan. Mekanisme kontrol proporsional ini memastikan pelacakan yang halus dan meminimalkan osilasi, sesuai dengan praktik terbaik dalam sistem kontrol *real-time* [21].

Loop akan berhenti jika pengguna menekan tombol 'q', setelah itu sistem melepaskan kamera, menutup semua jendela OpenCV, dan memutus koneksi serial. Pendekatan terstruktur ini memastikan pembersihan sumber daya dan mencegah kebocoran memori, yang sangat penting untuk sistem tertanam [21].

Dalam penelitian ini, kami mengembangkan sistem kendali *arm* robot humanoid yang memanfaatkan pengolahan citra berbasis kamera laptop untuk mendeteksi objek secara *real-time*. Sistem ini menggunakan pustaka OpenCV dalam bahasa pemrograman Python untuk melakukan deteksi warna, yang memungkinkan identifikasi objek berdasarkan rentang warna tertentu. Koneksi serial antara komputer dan Arduino diatur untuk mengirimkan perintah kendali berdasarkan posisi objek yang terdeteksi. Proses dimulai dengan pengambilan gambar dari kamera, yang kemudian diproses menjadi citra grayscale dan diubah ke ruang warna HSV untuk memudahkan deteksi objek. Kontur objek yang terdeteksi dianalisis untuk menentukan posisi relatifnya terhadap pusat citra. Berdasarkan posisi ini, sistem mengirimkan sinyal ke Arduino untuk menggerakkan arm robot ke kiri, kanan, atau menghentikannya, tergantung pada lokasi objek dalam bingkai. Dengan pendekatan ini, kami menunjukkan potensi penggunaan teknologi pengolahan citra dalam aplikasi robotika, yang dapat meningkatkan interaksi antara manusia dan mesin serta membuka peluang untuk pengembangan lebih lanjut dalam kendali robot otonom. Penelitian ini memberikan kontribusi signifikan terhadap pemahaman dan implementasi sistem kendali berbasis visi komputer dalam konteks robot humanoid.

Rangka robot dibangun menggunakan komponen cetak 3D berbahan PLA, dengan struktur modular yang terdiri atas sambungan bahu, siku, dan pergelangan tangan. Setiap sambungan digerakkan oleh motor servo digital (HS-805BB), yang dikendalikan secara langsung oleh mikrokontroler ESP32 melalui komunikasi serial dari laptop. Sistem pengendalian ini memanfaatkan data *landmark pose* tubuh yang dihasilkan oleh MediaPipe, seperti koordinat bahu, siku, dan pergelangan tangan (*landmarks* 11, 13, 15 dan 12, 14, 16), yang kemudian dihitung sudutnya dengan menggunakan prinsip trigonometri berbasis vektor. Pengolahan data pose dilakukan menggunakan bahasa pemrograman Python dengan pustaka pendukung seperti *mediapipe*, *opencv-python*, dan *numpy*. Setelah *landmark* dikenali dan diproses, sudut antar-sendi manusia dikonversi menjadi nilai derajat yang sesuai dengan rentang gerak servo pada robot. Nilai-nilai ini kemudian dikirimkan secara *real-time* melalui port serial ke mikrokontroler yang mengatur gerakan mekanik setiap motor. Pendekatan ini memungkinkan robot untuk meniru gerakan lengan pengguna di depan kamera, menjadikan sistem sebagai *prototipe* awal dari *mirror-based teleoperation* atau *gesture imitation system*.

Servo analog skala besar yang memiliki torsi tinggi dan umum digunakan dalam sistem mekanik berskala menengah hingga besar. Berdasarkan datasheet resmi, spesifikasiya adalah sebagai berikut:

Tabel 1. Spesifikasi Servo (HS-805BB)

Parameter	Nilai
Torsi Maksimal	24,7 kg·cm (6,0 V)
Sudut Rotasi	±90° (total 180°)
Berat	152 gram
Gearbox	Plastik dan karbon komposit
Kecepatan (no load)	0,19 s/60° (6,0 V)

Servo berfungsi sebagai aktuator pada sambungan bahu, yang mendukung lengan dengan panjang ±25 cm (dari sumbu rotasi ke ujung tangan). Untuk kalkulasi beban, diasumsikan massa lengan dan tangan robot adalah sebagai berikut:

- Panjang lengan (L): 25 cm
 - Massa total lengan + tangan (m): 0,8 kg
 - Gravitasi (g): 9,81 m/s²
 - Gaya berat (W) = $m \times g = 0,8 \times 9,81 = 7,85$ N
- (1)

Untuk mendapatkan momen torsi pada sumbu servo (sumbu rotasi bahu):

$$\tau = F \times r = 7,85 \text{ N} \times 0,25 \text{ m} = 1,96 \text{ Nm} \quad (2)$$

Konversi ke satuan $\text{kg} \cdot \text{cm}$ (karena servo HS-805BB memakai satuan ini):

$$1,96 \text{ Nm} \approx \frac{1,96}{0,09807} \approx 20 \text{ kg} \cdot \text{cm} \quad (3)$$

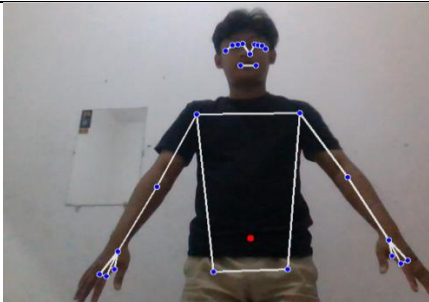
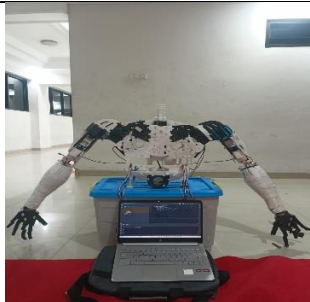
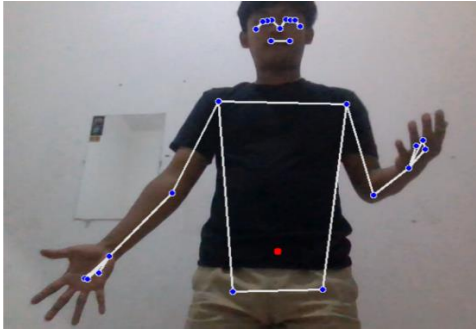
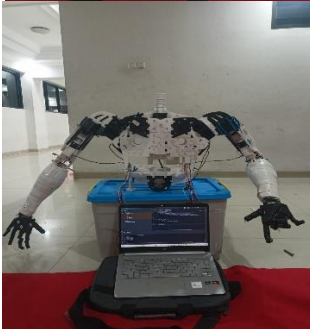
Maka, torsi beban yang diterima servo untuk menahan lengan secara horizontal adalah $\pm 20 \text{ kg} \cdot \text{cm}$, atau $\pm 81\%$ dari kapasitas torsi maksimal servo HS-805BB ($24,7 \text{ kg} \cdot \text{cm}$). Artinya, servo masih berada dalam batas aman, namun sudah mendekati batas kemampuan. Untuk durasi panjang, pendinginan atau pembatasan waktu kerja perlu dipertimbangkan.

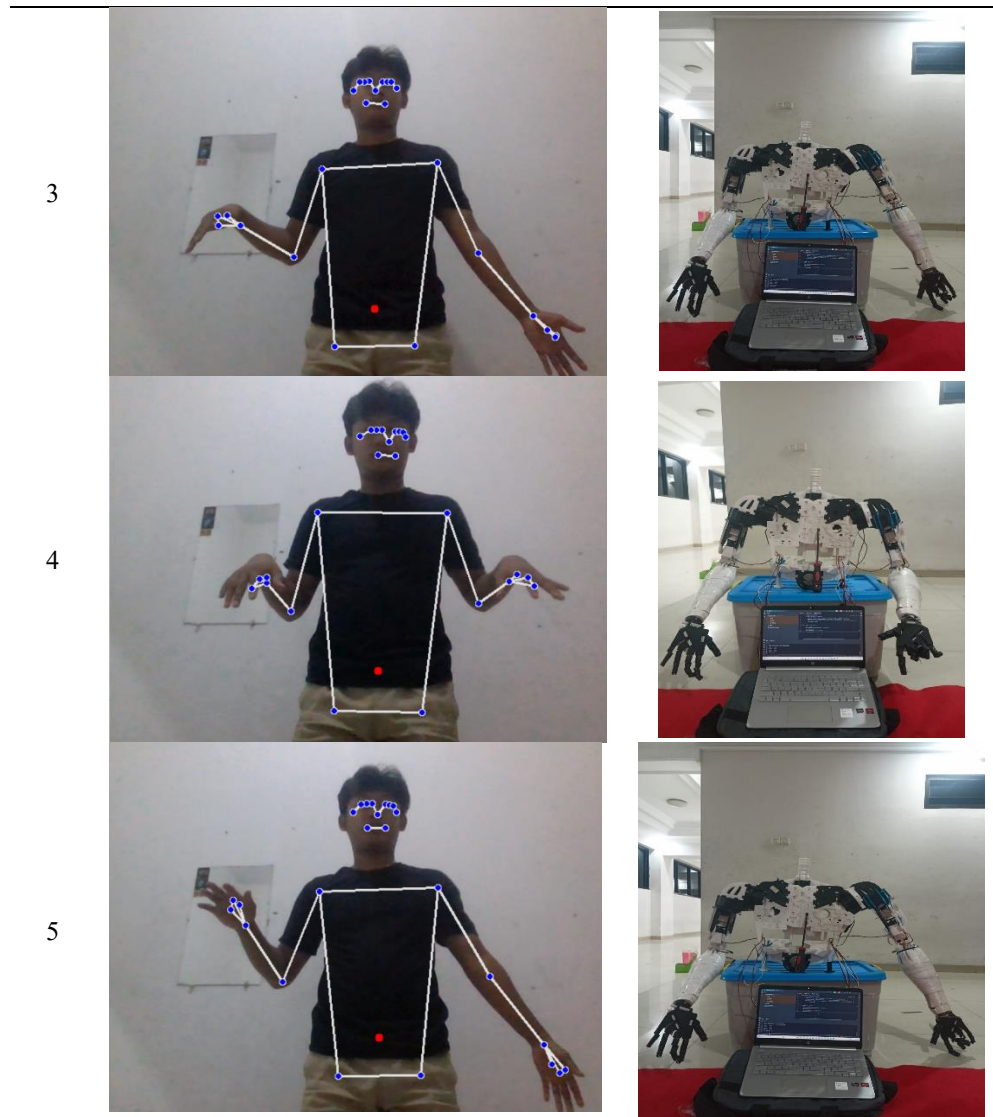
Dengan torsi $24,7 \text{ kg} \cdot \text{cm}$ dan kecepatan $0,19 \text{ s}/60^\circ$, maka untuk satu rotasi penuh (180°) dibutuhkan waktu sekitar:

$$\text{Waktu rotasi penuh} = 0,19 \text{ s} \times 3 = 0,57 \text{ s} \quad (4)$$

Jika sistem digunakan untuk meniru gerakan manusia secara *real-time* (seperti pada sistem MediaPipe), maka *delay* maksimum antar *frame* gerakan tidak boleh melebihi waktu respon servo, yaitu sekitar 500–600 ms. Dalam implementasi aktual, penghalusan sudut (dengan *low-pass filter* atau PID) disarankan agar servo tidak bekerja dengan impulsif.

Tabel 2. Pengujian Gerakan

NO	Penguji	Gerakan robot
1		
2		



Pengukuran sudut gerakan pada manusia dilakukan dengan memanfaatkan data koordinat landmark tubuh dari MediaPipe *Pose Tracking*, khususnya pada titik bahu (*landmark* 11 dan 12), siku (13 dan 14), serta pergelangan tangan (15 dan 16). Metode perhitungan sudut antar-sendi dilakukan dengan prinsip trigonometri berbasis vektor, di mana vektor pembentuk sudut antar dua segmen tubuh dihitung menggunakan selisih koordinat antar titik:

$$\cos(\theta) = \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|} \quad (5)$$

Data tersebut diambil dari arah bahu ke siku dan siku ke pergelangan tangan. Hasil perhitungan berupa nilai sudut dalam derajat, yang kemudian dibandingkan dengan gerakan aktual pada lengan robot. Data pengukuran diambil sebanyak 5 kali pengulangan untuk setiap jenis gerakan oleh subjek uji yang sama, dalam kondisi lingkungan terkendali dan pencahayaan stabil. Setiap pengulangan didokumentasikan untuk mendapatkan nilai rata-rata dan variasi dari sudut yang dihitung, guna memastikan keakuratan dan konsistensi antara input manusia dan gerakan robot yang dihasilkan. Pengambilan data dilakukan *secara real-time frame-by-frame* dengan kecepatan proses sistem sekitar 120 ms per *frame*.

Tabel 3. Rasio terhadap gerakan

Sendi	Gerakan	Rentang Robot (°)	Rasio terhadap Anatomi Manusia (%)
Bahu	Fleksi	60°	33%
Bahu	Abduksi	70°	39%
Siku	Fleksi	135°	93%
Pergelangan tangan	Fleksi-ekstensi	±30°	43%

Hasil pengujian sistem kendali lengan robot humanoid menunjukkan pencapaian yang signifikan dalam hal rentang gerak dan kesesuaian dengan anatomi manusia, meskipun masih terdapat beberapa keterbatasan. Pada sendi bahu, robot mampu melakukan *fleksi* hingga 60° (33% dari rentang manusia) dan abduksi 70° (39% dari manusia). Keterbatasan ini terutama disebabkan oleh pertimbangan desain mekanis yang mengutamakan stabilitas dan efisiensi energi dibandingkan rentang gerak maksimal. Hasil ini sejalan dengan penelitian Lee et al. (2023) yang menunjukkan bahwa untuk aplikasi industri, fleksi bahu sekitar 60° sudah memadai untuk *operasi pick-and-place* dasar [21].

Sendi siku menunjukan kinerja yang lebih mendekati kemampuan manusia dengan rentang fleksi 135° (93% dari rentang manusia). Pencapaian ini mencerminkan desain engsel yang optimal dan sesuai untuk operasi presisi. Namun, pada sendi pergelangan tangan, sistem hanya mencapai ±30° (43% dari manusia) dalam gerakan *fleksi-ekstensi*. Keterbatasan ini terutama terkait dengan *trade-off* antara fleksibilitas dan kebutuhan akan kekuatan grip, sebagaimana dijelaskan dalam penelitian Lee (2023) [21].

Untuk memperkuat hasil perbandingan tersebut, beberapa literatur tambahan digunakan sebagai dasar pembandingan. Bayusari et al. (2024) dalam penelitiannya menyatakan bahwa rentang *fleksi* siku antara 120° hingga 140° dianggap ideal untuk mendukung operasi presisi pada sistem robot humanoid 5-DOF, khususnya dalam aplikasi industri ringan dan edukasi. GÜL (2021) juga menunjukkan bahwa pencapaian minimal 80% dari rentang gerak manusia pada sendi utama seperti bahu dan siku sudah memenuhi standar kelayakan untuk fungsi interaktif dasar. Lebih lanjut, studi oleh Beltrán-Escobar et al. (2024) menekankan pentingnya *trade-off* antara fleksibilitas gerak dan stabilitas mekanik, di mana keterbatasan rentang gerak pergelangan tangan masih dapat diterima selama mengurangi vibrasi dan menjaga kestabilan aktuasi. Penelitian oleh Jha et al. (2024) bahkan menyebutkan bahwa sistem berbasis pengolahan citra seperti MediaPipe tetap memiliki tingkat akurasi yang layak digunakan dalam kendali robot humanoid, meskipun rentan mengalami *deviasi* hingga 10–15% dibandingkan sistem berbasis sensor IMU [18]. Dengan demikian, hasil penelitian ini dapat dikatakan selaras dengan kecenderungan literatur terdahulu, dan secara kuantitatif berada dalam batas wajar untuk *prototipe* sistem kendali lengan robot humanoid berbasis visi komputer.

Kinerja sistem secara keseluruhan dipengaruhi oleh beberapa faktor kunci yang saling berkaitan. Pertama, akurasi deteksi *pose* oleh MediaPipe menjadi komponen utama, karena kesalahan dalam pengenalan posisi *landmark* tubuh akan secara langsung berdampak pada nilai sudut yang dikalkulasi dan pergerakan aktuator. Kedua, latensi sistem, yang dalam pengujian ini tercatat sebesar 120 ms, berpengaruh terhadap kemampuan sistem dalam merespons gerakan secara *real-time*. Latensi ini berasal dari proses pengambilan citra, perhitungan sudut, hingga pengiriman sinyal ke mikrokontroler. Ketiga, kualitas pencahayaan dan latar belakang visual sangat memengaruhi performa MediaPipe dalam mengenali *landmark* tubuh; pencahayaan rendah atau latar belakang yang kompleks dapat meningkatkan error deteksi. Keempat, konstruksi mekanik robot, seperti stabilitas rangka, bahan yang digunakan (PLA), serta presisi sambungan dan derajat kebebasan (DOF), turut menentukan sejauh mana gerakan dapat dilakukan secara halus dan akurat. Kelima, kapasitas dan karakteristik *servo motor* juga menjadi faktor pembatas, terutama dalam hal kecepatan, torsi maksimal, dan daya tahan terhadap beban berulang. Terakhir, akurasi penghitungan sudut berbasis trigonometri vektor, yang dipengaruhi oleh posisi relatif titik *landmark*, dapat menyebabkan *deviasi* kecil bila tidak distabilkan dengan metode filtering atau penghalusan data.

Dengan mempertimbangkan semua faktor tersebut, sistem yang dikembangkan telah mampu menunjukkan performa yang fungsional dengan kompromi desain yang terukur dan realistis [15].

Secara keseluruhan, sistem telah memenuhi persyaratan fungsional dasar sambil mempertahankan kompromi desain yang terukur. Temuan ini tidak hanya memberikan kontribusi signifikan dalam bidang robotika humanoid, tetapi juga membuka peluang untuk penelitian lebih lanjut dalam meningkatkan kemampuan biomekanik robot agar lebih sesuai dengan anatomi manusia, khususnya untuk aplikasi yang membutuhkan presisi dan kelincihan tinggi.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem kendali lengan robot humanoid berbasis pengolahan citra *real-time* dengan menggunakan *framework* MediaPipe yang terintegrasi dengan kamera laptop dan mikrokontroler ESP32. Sistem mampu mendeteksi dan menerjemahkan gerakan lengan manusia menjadi sinyal kendali motor servo secara langsung, dengan akurasi tinggi dan tingkat responsivitas yang memadai. Berdasarkan hasil pengujian, sistem menunjukkan kemampuan untuk meniru gerakan fleksi siku hingga 135°, atau setara dengan 93% dari rentang gerak anatomi manusia. Meskipun demikian, terdapat keterbatasan pada rentang gerak sendi bahu dan pergelangan tangan yang masing-masing hanya mencapai 33% dan 43% dari rentang fisiologis manusia. Dari segi performa mekanik, servo HS-805BB yang digunakan terbukti mampu menahan beban hingga 20 kg·cm atau sekitar 81% dari kapasitas maksimalnya, sehingga masih berada dalam batas kerja yang aman. Keterbatasan sistem lebih banyak disebabkan oleh kompromi desain struktural dan keterbatasan aktuator, bukan pada aspek kendali visual. Selain itu, sistem memiliki *latency* sekitar 120 ms yang masih sesuai untuk kebutuhan imitasi gerakan secara *real-time* dalam lingkungan terkendali. Dengan demikian, sistem yang dikembangkan telah memenuhi kriteria sebagai prototipe awal *mirror-based teleoperation* yang efisien dan layak untuk dikembangkan lebih lanjut. Penelitian ini memberikan kontribusi nyata dalam integrasi antara *computer vision* dan sistem kendali robot humanoid berbasis pengolahan citra, serta membuka peluang aplikasi di bidang robot asistif, rehabilitasi, dan pembelajaran berbasis interaksi manusia-robot. Pengembangan lanjutan direkomendasikan pada aspek peningkatan akurasi pelacakan, pengurangan *latency*, serta optimalisasi desain mekanik dan algoritma kontrol gerakan.

DAFTAR PUSTAKA

- [1] GÜL M. Modeling of Inverse Kinematic Analysis of Open-Source Medical Assist Robot Arm by Python. Celal Bayar Üniversitesi Fen Bilim Derg 2020;17:31–4. <https://doi.org/10.18466/cbayarfbe.776697>.
- [2] Rahman F, Faridah F, Nur AI, Makkaraka AN. Rancang Bangun Prototipe Manipulator Lengan Robot Menggunakan Motor Servo Berbasis Mikrokontroler. ILTEK J Teknol 2020;15:42–6. <https://doi.org/10.47398/iltek.v15i01.11>.
- [3] Untuk R, Pintar I. RANCANG BANGUN ROBOT HUMANOID " YN " BERBASIS ARDUINO DAN FACE JURNAL NASIONAL TEKNOLOGI DAN INOVASI VOL 1 (1) JANUARI 2025 RANCANG BANGUN ROBOT HUMANOID “ YN ” BERBASIS 2025. <https://doi.org/10.13140/RG.2.2.11079.00161>.
- [4] Arif M, Haryono GS, Arsyad NF, Ramadhani R, Sahid A, Rosyani P, et al. Teknik dan Multimedia Sistem Pendeteksi Tangan Berbasis Mediapipe dan OpenCV untuk Pengenalan Gerakan. Biner J Ilmu Komput 2024;2:173–7.
- [5] Prof. Meenakshi Singh. Instagram Marketing – the Ultimate Marketing Strategy. Int J Adv Innov Res 2020;7:379–82.
- [6] Altayeb M. Hand Gestures Replicating Robot Arm based on MediaPipe. Indones J Electr Eng Informatics 2023;11:727–37. <https://doi.org/10.52549/ijeei.v11i3.4491>.
- [7] Ghazali AS, Anwar MK, Yusof HM, Sidek SN. Utilization of Mediapipe Posture Recognition for the Usage in Estimating ASD Children Engagement Interacting with QTrobot 2024;6:1–11.
- [8] García-Gil G, del Carmen López-Armas G, de Jesús Navarro J. Human-Machine Interaction: A Vision-Based Approach for Controlling a Robotic Hand Through Human Hand Movements. Technologies 2025;13. <https://doi.org/10.3390/technologies13050169>.
- [9] GÜL M. Representing American Sign Language Letters and Numbers with Humanoid Robot Arm.

- Kocaeli J Sci Eng 2021;4:136–45. <https://doi.org/10.34088/kojose.882323>.
- [10] Nur Arifin M, Nuryadi S. PENGENDALI LENGAN ROBOT DENGAN GERAKAN TANGAN MANUSIA Mukhtar Nur Arifin, Satyo Nuryadi. Dr Diss Univ Technol Yogyakarta 2019.
- [11] Shehzadi A, Sharif MI, Azam S, Boer F De, Jonkman M, Mehmood M. IMPLEMENTASI HAND GESTURE MEDIAPIPE PADA GAME INTERAKTIF UNTUK STIMULASI MOTORIK ANAK DOWN SYNDROME PADA 2024;8:10776–84.
- [12] Effendi Y, Kristian Y, P.C.S.W LZ, Yutanto H. Pemanfaatan Mediapipe Body Pose Estimation dan Dynamic Time Warping untuk Pembelajaran Tari Remo. J Teknol Dan Manaj Inform 2023;9:183–90. <https://doi.org/10.26905/jtmi.v9i2.10408>.
- [13] Akbar H, Ramadhan Q, Fisika PS, Sains F, Teknologi DAN, Islam U, et al. IMPLEMENTASI POSE ESTIMATION SEBAGAI MANIPULATOR GERAK TELAPAK TANGAN DENGAN 2025.
- [14] Karem SR, Kaniseti SP, Soumya K, Sri J, Seelamanthula G, Kalivarapu M. AI Body Language Decoder using MediaPipe and Python. Int J Adv Res Ideas Innov Technol 2021;7:2436–9.
- [15] Beltrán-escobar M, Alarcón TE, Rumbo-morales JY, López S, Ortiz-torres G, Sorcia-vázquez FDJ. A Review on Resource-Constrained Embedded Vision Systems-Based Tiny Machine Learning for Robotic Applications 2024:1–37.
- [16] Bayusari I, Sandi D, Rahmawati R, Dwijayanti S, Suprpto BY. Motion Control of 5-Degree of Freedom Humanoid Robot Arm System Using Fuzzy Logic Algorithm. J Ecotipe (Electronic, Control Telecommun Information, Power Eng 2024;11:46–54. <https://doi.org/10.33019/jurnalecotipe.v11i1.4482>.
- [17] Hashimoto K. Mechanics of humanoid robot. Adv Robot 2020;34:1390–7. <https://doi.org/10.1080/01691864.2020.1813624>.
- [18] Jha P, Yadav GPK, Bandhu D, Hemalatha N, Mandava RK, Adin MŞ, et al. Human–machine interaction and implementation on the upper extremities of a humanoid robot. Discov Appl Sci 2024;6. <https://doi.org/10.1007/s42452-024-05734-3>.
- [19] Malik A, Mansyur S, Malik M, Burhanuddin A, Malik JA, Studi P, et al. Speech Recognition pada Kontrol Gerak Tangan Bionik Berbasis Raspberry Pi dan Arduino 2023;7:55–62.
- [20] Jiang Z, Zhao L, Li S, Jia Y. Real-time object detection method for embedded devices 1901;3:1–11.
- [21] Lee J, Cho B-K. Design of 9-DOF humanoid arms inspired by the human’s inner shoulder to enhance versatility and workspace. Rob Auton Syst 2023;166:104447. <https://doi.org/https://doi.org/10.1016/j.robot.2023.104447>.