Pose Estimation-Driven Control of Humanoid Upper Arms for Human Motion Mimicry

Omar Salem, A. Abdellatif *, and Mostafa R. A. Atia

Mechanical Engineering Department, Arab Academy for Science Technology and Maritime Transport, Cairo, Egypt
Email: omaradeldh3@gmail.com (O.S.); a_abdellatif@aast.edu (A.A.); mrostom1@aast.edu (M.R.A.A.)
*Corresponding author

***Abstract*—In the modern era, Robotics has become an essential part of modern life, enhancing efficiency and precision in manufacturing and supporting diagnosis, prediction, and surgery in medicine. This paper presents a real-time teleoperation framework that maps human upper-body motion, captured by a single RGB-D camera, to a dual-arm upper humanoid robot designed with low-cost servos. The system employs MediaPipe-based pose estimation and a torso-anchored coordinate transformation to achieve operator-centric retargeting that is robust to variations in camera placement and subject geometry. To suppress tremor and sensor noise, a constant-velocity Kalman filter combined with an adaptive dead-zone is applied to wrist trajectories, ensuring smooth motion while maintaining responsiveness. A min-max scaling function with saturation enforces safe workspace mapping, while joint commands are computed using damped-least-squares inverse kinematics with joint-limit and self-collision checks. The execution layer incorporates incremental speed-aware stepping to emulate continuous trajectories on servo actuators. Experimental results demonstrate accurate static pose reproduction, robust dynamic path following, and zero joint-limit violations, achieving an average wrist-tracking Root Mean Square Error (RMSE) of 12.4 mm and median end-to-end latency of 86 ms. The platform is reproducible, cost-effective, and adaptable for applications in education, rehabilitation, and human–robot collaboration.**

***Keywords*—humanoid robotics, teleoperation, human-robot interaction, robot kinematics, motion capture system, image processing, signal processing, Kalman filters**

I. Introduction

The human hand is regarded as a model of dexterity and precision, performing tasks that range from lifting heavy objects to manipulating delicate items [1]. Replicating such versatility has long posed a significant engineering challenge, driving advances in robotics, prosthetics, Virtual Reality (VR), and Human-Computer Interaction (HCI) [2]. In robotics, human-like hand motion is replicated to enable machines to conduct intricate operations with high precision, particularly valuable in industrial applications where accuracy and adaptability are essential [3].

For users with limb differences, prosthetic hands designed to imitate natural movement have been shown to restore function and improve quality of life [4]. Their development relies on the integration of biomechanics and intelligent control systems, with realistic motion replication remaining a major focus [5]. In VR, accurate hand tracking enhances immersion by ensuring seamless correspondence between real and virtual movements, transforming training and educational experiences [6]. Within HCI, interaction has been increasingly driven by gesture-based control, supported by motion-tracking technologies capable of converting hand gestures into digital commands [7]. These advancements continue to expand applications in smart environments, interactive systems, and professional contexts, redefining the way humans interact with digital systems [8].

Recent advancements in Human Pose Estimation (HPE) have significantly improved both accuracy and computational efficiency. Comparative analyses have indicated that MoveNet is particularly effective for real-time detection, making it suitable for dynamic applications such as sports analysis and medical assistance [9]. Earlier CNN-based models and the OpenPose framework established key architectural benchmarks using datasets like COCO and MPII [10, 11]. Later enhancements in the MediaPipe framework strengthened 3D posture estimation by mitigating depth and joint localization errors, improving reliability across users [12]. The combination of MediaPipe's BlazePose with Kinect data was found to enhance action recognition through advanced visualization techniques [13].

Furthermore, MediaPipe outperformed Kinect in humanoid robot control, offering more precise real-time tracking of 3D joint positions [14]. Collectively, these developments demonstrate the rapid evolution of HPE methods and their increasing potential for accurate motion interpretation in robotics and related domains [15, 16].

Recent progress has extended the application of MediaPipe beyond posture tracking to include force estimation and fall detection. In fall detection, a vision-based model achieved 95.84% accuracy by analyzing descent speed, height reduction, and bounding box variation, outperforming traditional sensor-based

Manuscript received October 11, 2025; revised November 20, 2025; accepted January 5, 2026; published April 23, 2026.

doi: 10.18178/ijmerr.15.2.174-186

systems with fewer false alarms and greater reliability [17, 18]. Concurrently, the ForcePose framework combined skeletal tracking and object detection to estimate human–object interaction forces. Tests on 850 annotated videos yielded mean errors of 5.83 N in magnitude and 7.4° in direction, marking a 27.5% improvement over conventional techniques [19, 20].

These results underscore the growing precision of vision-based methods in quantifying human motion and interaction. Pose estimation systems such as MediaPipe BlazePose and PoseNet have also been widely employed in health and fitness applications, enabling real-time tracking of posture and joint movement for use in injury prevention and performance assessment [12, 20, 21].

Further developments introduced precise yoga pose detection using skeletonized images, as well as innovative 3D approaches that combined 2D estimation with humanoid modeling. These methods were particularly suited for applications such as real-time elderly care, where accuracy and robustness are critical [22, 23]. Research during this period also emphasized the importance of context-dependent accuracy in physical therapy and identified performance limitations—along with potential improvements—for widely used platforms such as OpenPose and MediaPipe [24]. Building on these advances, highly accurate home rehabilitation systems were developed using MediaPipe. Such systems demonstrated the ability to classify upper-limb movements reliably, offering new opportunities to enhance both health monitoring and fitness training [25, 26].

With the rapid advancement of computer vision and HPE, recent studies have demonstrated remarkable progress in real-time detection and analysis. Powered by deep learning, these systems are reshaping domains such as gaming, entertainment, and sports by enabling instantaneous performance monitoring and highly accurate motion insights [27]. One notable example is the evaluation of Microsoft's MediaPipe framework on the Martial Arts, Dancing, and Sports (MADS) dataset. The framework achieved 3D pose estimation accuracy consistently in the range of 100–200 mm, highlighting its suitability for demanding applications such as advanced sports performance analysis, seamless human–machine collaboration, and enhanced surveillance systems [28].

Beyond datasets, MediaPipe has been extended to diverse media inputs, including internet images and live video streams. Such efforts have demonstrated their ability not only to estimate joint angles for everyday activities but also to support the development of web-based interfaces aimed at substituting physical trainers and promoting health and sports activity tracking. Collectively, these advancements illustrate the transformative potential of MediaPipe and similar paradigms in capturing and translating human movement across a wide range of real-world applications [29].

This study was conducted to design and validate a real-time humanoid arm mimicry system that enables safe, affordable, and intuitive human–robot interaction in high-risk environments. The MediaPipe Pose framework, combined with an Astra Pro RGB-D camera, was employed to offer a low-cost and flexible alternative to traditional motion-capture systems, allowing camera placement anywhere within the operator's workspace. A torso-anchored retargeting method was implemented to maintain robustness against camera-position changes, while a Kalman-filter-based motion pipeline with a spatial dead zone was applied to minimize jitter and tremors. Proportional scaling, integrated with damped least-squares inverse kinematics and joint-limit constraints, was used to ensure smooth and collision-free motion. The overall framework unified pose detection, coordinate transformation, and dual-arm inverse kinematics with real-time monitoring, resulting in a scalable and environment-independent teleoperation system suitable for educational, rehabilitative, and collaborative applications.

Unlike previous systems that rely on multi-camera setups, wearable markers, or computationally intensive deep-learning models, the proposed framework introduces a single-camera, marker-less humanoid mimicry system capable of real-time dual-arm teleoperation. Through the integration of MediaPipe Pose-based tracking, Kalman-filtered spatial mapping, and inverse kinematics control, the system achieves millimeter-scale precision and low-latency motion replication. This approach provides a low-cost, robust, and scalable solution for intuitive human-robot interaction and rehabilitation robotics.

The remainder of this paper is organized as follows: Section II describes the research methodology and the main components of the humanoid mimicry system. Section III presents experimental work and validation procedures. Section IV provides a detailed discussion of the results obtained. Finally, Section V concludes the paper and outlines potential directions for future development.

II. Methodology

A. Mechanical and Electrical Design

For the assembly of the current robotic hand (Table I), a combination of locally sourced materials and one imported component—the camera, procured from China.

As shown in Fig. 1, the design focused on the humanoid upper body, emphasizing actuator placement within the torso. Servomotors were arranged at the acromioclavicular joints (1 and 5), shoulder joints (2 and 6), elbow joints (3 and 7), and wrist joints (4 and 8). This configuration was optimized to replicate natural upper-limb motion, enabling precise and coordinated movements suitable for advanced humanoid robotic applications.

The robotic arm's electrical system was configured to ensure stable and efficient power delivery across all components. The AC supply was converted to 12 V DC through an AC-DC converter, then stepped down to 10 V DC by a buck converter to meet the servo motors' requirements. All servos were connected in parallel for consistent power distribution. Motion control was managed by an Arduino Mega generating PWM signals, Fig. 2, while high-level commands were transmitted from a Python-based interface via serial communication. This configuration enabled precise and adaptive motion control, improving the arm's responsiveness and accuracy in complex tasks [30].

TABLE I. ILLUSTRATED THE ESSENTIAL HARDWARE COMPONENTS USED TO BUILD AND OPERATE THE DUAL ROBOTIC ARMS OF THE HUMANOID SYSTEM

Component Name	Quantity	Key Function Summary
Power Supply Adapter	1	Provides stable 12 V DC power for all 8 servo motors.
Power Supply Step Down Module	1	Regulates 12V down to a stable 8–10 V to ensure consistent servo performance and prevent spikes.
Digital Metal Gear Servo (35 kg.cm, 180°)	2	Controls primary shoulder movement (Y-Z plane) for overall arm positioning.
25kg.cm Digital Servo (180°)	4	Controls arm extension (elbow) and wrist/hand rotation for precise positioning (Two per arm).
FEETECH FT5835M Digital Servo	2	Enables critical 0–90° upper arm rotation to transition movement between vertical and horizontal planes (side-to-side reach).
ARDUINO MEGA	1	Slave controller that receives commands and generates PWM signals to coordinate all arm movements.
3D Print of Humanoid Upper Body	1	Structural shell housing arms, servos, Arduino, and power/communication circuit boards.
[Sensor/Camera System]	1	Captures human movement (RGB and Depth) and translates joint positions into robot coordinates for mimicry.

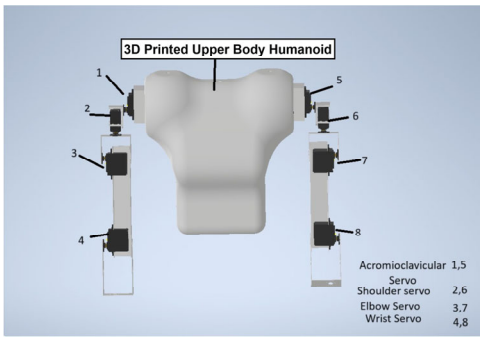


Fig. 1. Upper humanoid robot layout.

B. *Vision-Based Pose Estimation*

To achieve 3D perception, the system utilizes an Astra Pro RGB-D camera based on structured-light technology. By projecting an IR grid and analyzing the pattern's distortion through triangulation, the camera calculates pixel-wise depth from a single vantage point (Fig. 3). The resulting high-fidelity depth map transforms 2D visual data into a 3D spatial representation, providing the necessary data structure for MediaPipe pose estimation.

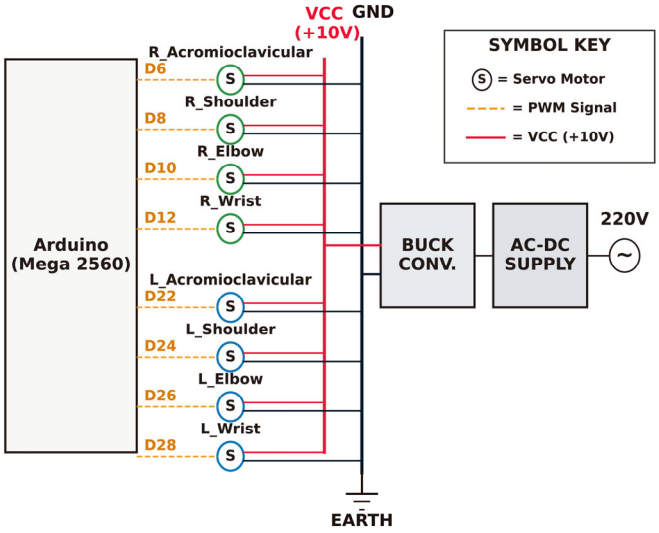


Fig. 2. Electrical wiring diagram.

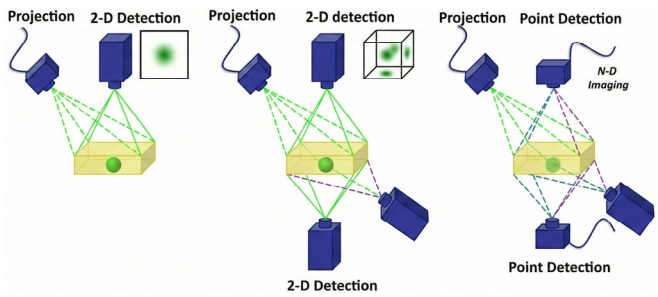


Fig. 3. Comparison of imaging-based detection schemes for spatial perception. (a) 2-D single-view projection; (b) stereo vision with triangulated 3-D localization; and (c) direct depth-enabled point detection [31].

C. Coordinate Transformation and Scaling

To convert raw pixel data into a 3D representation, the system was calibrated using the intrinsic parameters of the RGB-D camera—its digital fingerprint. These parameters defined how the camera interpreted its surroundings and enabled the transformation of 2D image coordinates into 3D spatial positions [32]. The key parameters included the focal lengths $(f_x, f_y)$and the optical center (c_x, c_y). The focal lengths, derived from the field of view and resolution, represented the effective zoom levels, while the optical center indicated the intersection of the optical axis with the image sensor, compensating for lens distortion and manufacturing tolerances. Once calibration was completed, each 2D pixel coordinate (u, v) with an associated depth value D was projected into real-world 3D space using the following Eq. (1):

$$Z = D, X = \frac{(u-c_x)\times Z}{f_x}, Y = \frac{(v-c_y)\times Z}{f_y} \quad (1)$$

After the 3D coordinates were obtained, a min-max scaling procedure was applied to proportionally map human hand movements into the robot's limited workspace. This linear interpolation approach, commonly referred to as range mapping, is widely used in science and engineering to translate values between domains [33]. The general form is shown in Eq. (2):

$$y = y_{min} + (y_{max} - y_{min}) \times \frac{(x-x_{min})}{(x_{max}-x_{min})} \quad (2)$$

Here, x represents the input value within the original range $[x_{min}, x_{max}]$, and y denotes the corresponding mapped value within the target range $[y_{min}, y_{max}]$.

D. Kinematic Modeling

The process begins once the desired coordinates of the robot's hand are determined. The system then solves the inverse kinematics problem, calculating the exact joint angles required to reach the target position. Before motion is initiated, these angles are validated to ensure they lie within the robot's mechanical limits. If the position is feasible, the computed angles are transmitted to the servo motors for execution; otherwise, visual feedback is provided to alert the user that the target is unreachable [34].

Denavit-Hartenberg (DH) convention was employed to define spatial relationships between successive links. Each joint's position and orientation relative to the previous one were represented by a 4×4 transformation matrix derived from the four geometric parameters as shown in Eq. (3).

$$T_i^{i-1} = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \cos\alpha_i & \sin\theta_i \sin\alpha_i & a_i \cos\theta_i \\ \sin\theta_i & \cos\theta_i \cos\alpha_i & -\cos\theta_i \sin\alpha_i & a_i \sin\theta_i \\ 0 & \sin\alpha_i & \cos\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3)$$

E. Speed Control

Speed control was achieved by calculating the real-time velocity of the user's hand and converting it into a discrete stepped value. The 3D wrist position was continuously tracked across consecutive frames, and a Kalman filter was applied to the raw position data to suppress noise and eliminate jitter [35]. The linear velocity in millimeters per second (mm/s) was then computed from the displacement of the filtered hand position divided by the elapsed time.

This continuous speed was subsequently converted into a discrete step ranging from 1 to 15 using a scaling function. The computed velocity was first clamped to a maximum threshold (e.g., 500 mm/s) to reject unrealistic motion, then linearly mapped across the defined range, where 0 mm/s corresponded to step 1 and the maximum velocity to step 15. The resulting "speed step" was employed to control the robot's motor velocity, enabling the system to replicate both the position and dynamic motion of the human operator (Fig. 4).

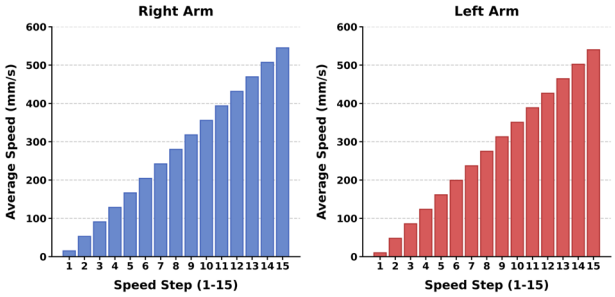


Fig. 4. Average speed (mm/s) vs Scaled speed step.

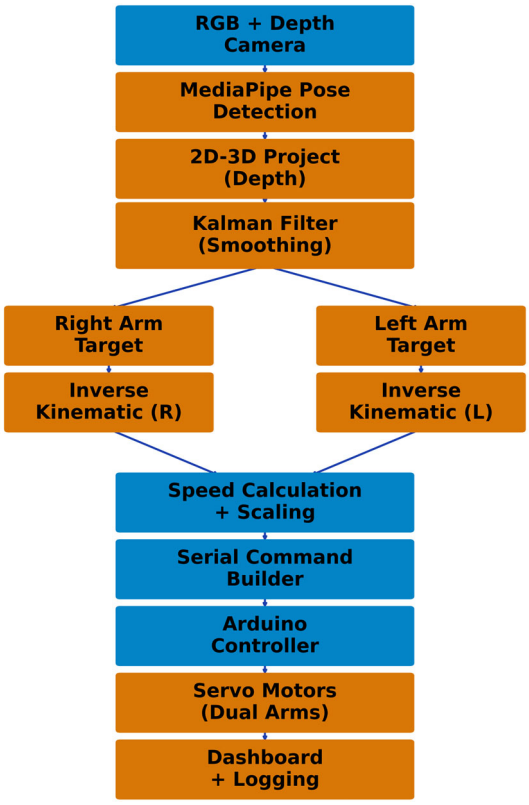


Fig. 5. System workflow.

F. RGB-D Pose to Dual-Arm Robot Control

The flowchart in Fig. 5 summarizes the complete methodology pipeline of the humanoid mimicry system.

Human motion was captured by an RGB-D camera and processed through MediaPipe for pose detection, 3D projection, and Kalman filter–based smoothing with velocity estimation. Target positions for both arms were computed and solved using inverse kinematics, after which speed scaling was applied. The resulting motion commands were then transmitted to an Arduino controller to drive the dual-arm servomotors. System performance was continuously monitored through a real-time dashboard and logged for analysis.

III. Experimental Work

The implementation of the humanoid mimicry system is illustrated in Fig. 6. The figure shows the fully assembled 3D-printed humanoid torso, dual robotic arms, control laptop, and RGB-D camera operating as an integrated unit. Preliminary evaluations of mechanical alignment and motion confirmed correct assembly and system readiness for extended testing.

The laptop's dashboard served as the primary control interface, providing live monitoring of both arms' performance, including current and target joint angles, positional data, and servo velocities. The dashboard also displayed the real-time video feed of the human operator, whose movements were accurately mirrored by the robotic arms. This synchronization demonstrated the system's complete functionality—from human motion capture to precise robotic execution—validating the full control pipeline.

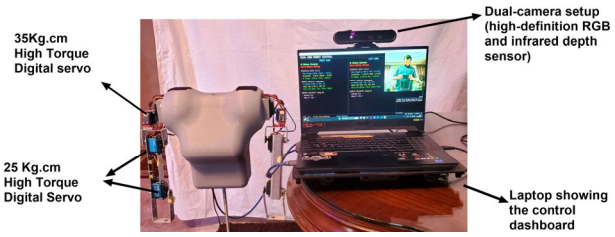


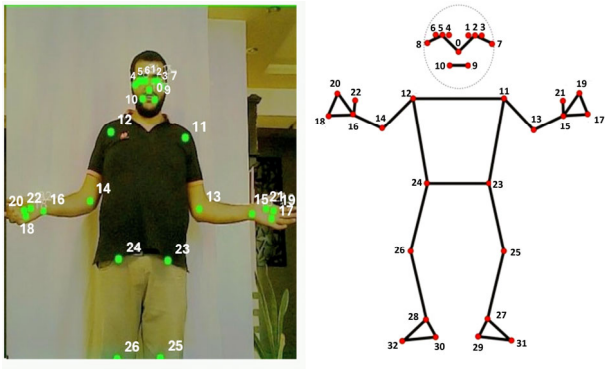
Fig. 6. Fully assembled system.

The experimental configuration was centered around an Astra Pro series camera, which was securely mounted on a vertical stand at a fixed height of 1500 mm. The system operator was positioned opposite the sensor at a working distance ranging between 1400 mm and 1600 mm. To ensure optimal sensor performance, the lighting environment was carefully controlled using indirect interior illumination to eliminate the risk of sensor saturation from direct exposure; furthermore, to evaluate the system's adaptability, distinct robustness tests were executed under natural sunlight conditions. Throughout these trials, data acquisition was maintained at a consistent sampling rate of 30 frames per second (fps).

In terms of calibration and accuracy, the Astra Pro hardware inherently possesses a factory-rated margin of error of ± 3 mm at 2000 mm. To improve upon this baseline and achieve higher fidelity, a software-based post-processing stage was introduced. This involved the implementation of a k-Nearest Neighbors (k-NN) algorithm, which was utilized to calculate the average of the depth readings. This computational step effectively normalized the depth field, smoothing out irregularities to enhance the overall precision of the system.

After the hardware configuration was validated, the robot's RGB camera was activated to generate a real-time video stream of the environment. This stream was processed through MediaPipe's pose detection algorithm, which continuously identified human subjects and extracted key anatomical landmarks such as the shoulders, elbows, and hands to form a skeletal model of the upper body.

As shown in Fig. 7, high detection accuracy was achieved, with real-time landmark identification (Fig. 7(a)) closely matching the reference points defined in the MediaPipe Library (Fig. 7(b)). This strong correspondence ensured the generation of a precise digital skeleton, providing a dependable foundation for motion tracking and robotic mimicry.



(a) (b)

Fig. 7. Visualization of anatomical reference points. (a) Landmarks are identified by the system in real time; (b) Comparison with official reference points from the MediaPipe Library.

When a person was detected, the robot's RGB-D camera was activated to perform two tasks simultaneously. First, it combined color and depth data to track the arm's two-dimensional (X, Y) trajectory through space, as illustrated in Fig. 8(a)–(c). At the same time, it generated a live, color-coded depth map to visualize proximity: blue indicated that the hand was near, yellow represented a neutral reference point (such as the torso), and red signified that the hand was farther away. This visualization was consistent across all axes—horizontal (X, Fig. 8(d)), forward/backward (Z, Fig. 8(e)), and vertical (Y, Fig. 8(f)). Notably, whenever the hand aligned with the chest, the map reliably returned to yellow, confirming that the depth visualization was independent of frame position.

To ensure that the robot's responses were smooth and intentional, the incoming 3D coordinates from three key markers, the midpoint between the shoulders (landmarks 11 and 12), and the two wrists (landmarks 15 and 16), were processed through a two-stage filtering pipeline. In the first stage, an Exponential Moving Average (EMA), parameterized by $\alpha = 0.25$, was applied to reduce high-frequency noise and smooth the trajectories while preserving responsiveness.

In the second stage, a spatial dead zone was enforced, discarding minor involuntary movements by issuing a new

command only if the filtered target moved more than 8.0 mm from its previous position. This combined strategy of EMA filtering and dead-zone enforcement proved essential for producing motion that was both smooth and purposeful. By suppressing jitter and unintentional tremors, the system ensured that the robot's actions consistently reflected the operator's intended gestures.

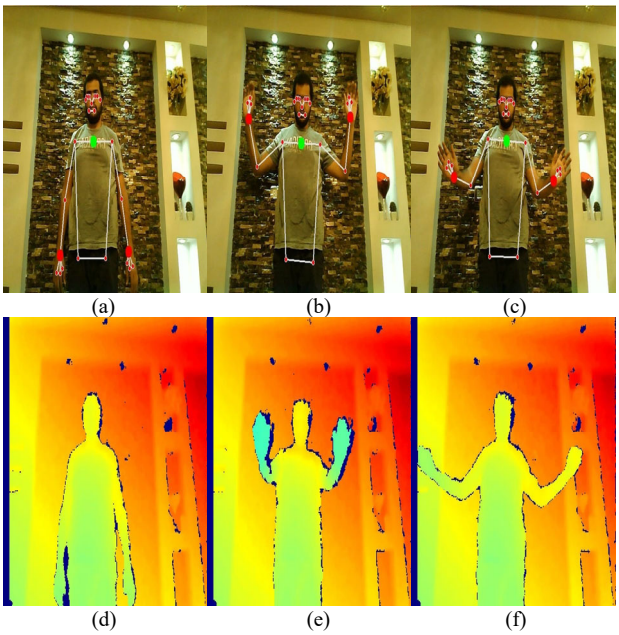


Fig. 8. Human pose and depth perception results for three representative postures. (a) RGB image showing detected skeletal key points for movement in the Y-axis; (b) RGB image showing skeletal tracking for movement in the Z-axis; (c) RGB image showing skeletal key points for movement in the X-axis; (d) depth-based silhouette illustrating the subject's spatial profile relative to the sensor for the Y-axis posture; (e) depth-based silhouette for the Z-axis posture; and (f) depth-based silhouette for the X-axis posture.

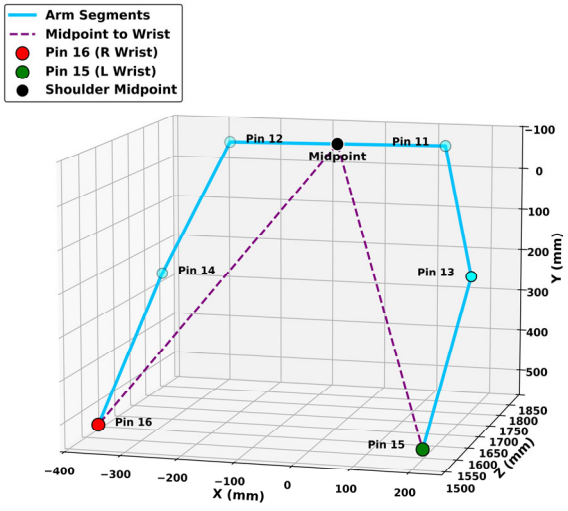


Fig. 9. Resultant poses (3D + data) for case A.

The current real-time video processing and landmark extraction confirmed the system's ability to accurately capture and analyze human pose data, as illustrated in Figs. 9–11. These figures present the system's direct output after applying the MediaPipe pose algorithm to 3 different human poses. The 3D visualization reconstructed the upper body by plotting spatial coordinates of key anatomical landmarks—including the shoulders (Pins 11, 12), elbows (Pins 13, 14), and wrists (Pins 15, 16)—and connecting them to form a digital skeleton.

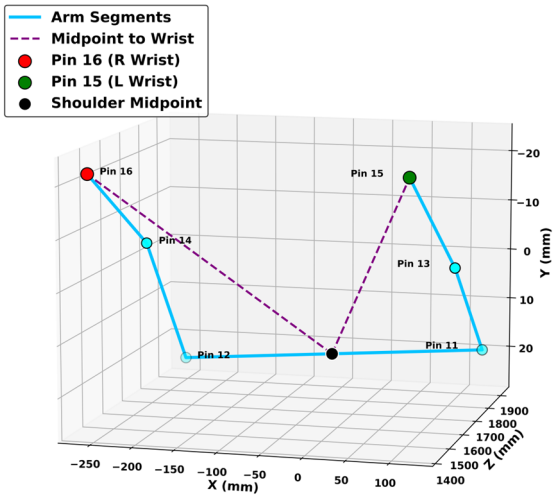


Fig. 10. Resultant poses (3D + data) for case B.

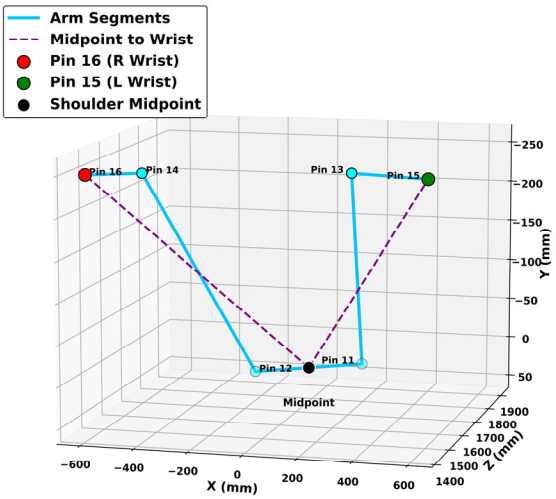


Fig. 11. Resultant poses (3D + data) for case C.

Tables II–IV list the raw X, Y, and Z coordinates of each landmark in millimeters, while additional computations were performed to derive secondary features such as the "Shoulder Midpoint" (green dot) and the vectors from this point to both wrists (pink dashed lines). This conversion of visual data into a structured numerical dataset provided precise spatial information essential for robotic mimicry.

TABLE II. THE RAW DATA AND CALCULATED DIFFERENCES IN CASE A

Landmarks	Case A		
	X (mm)	Y (mm)	Z (mm)
11 (L shoulder)	122.00	2183.00	45.77
13 (L Elbow)	214.02	2077.00	−315.72
15 (L Wrist)	262.95	1981.00	−619.57
12 (R shoulder)	−266.32	1794.00	68.96
14 (R Elbow)	−372.01	1775.00	−248.11
16 (R Wrist)	−402.71	1647.00	−512.23
Shoulder Midpoint	−72.16	1988.50	57.37
Diff Mid-to-L Wrist	−335.11	7.50	6.76.93
Diff Mid-to-R Wrist	330.55	341.50	569.59

TABLE III. THE RAW DATA AND CALCULATED DIFFERENCES IN CASE B

Landmarks	Case B		
	X (mm)	Y (mm)	Z (mm)
11 (L shoulder)	111.76	1882.00	46.04
13 (L Elbow)	136.76	1263.00	−101.51
15 (L Wrist)	133,84	1236.00	−32.39
12 (R shoulder)	−251.63	1757.00	46.05
14 (R Elbow)	−256.32	1223.00	−74.79
16 (R Wrist)	−248.54	1206.00	−29.50
Shoulder Midpoint	−69.94	1819.50	46.04
Diff Mid-to-L Wrist	−203.77	583.50	78.44
Diff Mid-to-R Wrist	178.61	613.50	75.54

A two-step process was implemented to convert visually captured human poses into a structured, robot-centric coordinate system using the measured data. Fig. 12 illustrates the initial raw data captured by the system. The 3D plot reconstructs the user's upper body by linking key anatomical landmarks, while Table V lists their absolute 3D coordinates—for example, the left shoulder (Landmark 11) at (X: 99.19, Y: −99.23, Z: 1721.00) and the right wrist (Landmark 16) at (X: −360.02, Y: 440.49, Z: 1483.00).

TABLE IV. THE RAW DATA AND CALCULATED DIFFERENCES IN CASE C

Landmarks	Case C		
	X (mm)	Y (mm)	Z (mm)
11 (L shoulder)	89.27	1893.00	39.69
13 (L Elbow)	209.11	1842.00	276.78
15 (L Wrist)	486.06	1784.00	-280.53
12 (R shoulder)	−248.98	1813.00	53.85
14 (R Elbow)	−415.19	1748.00	−244.33
16 (R Wrist)	−642.14	1578.00	−256.41
Shoulder Midpoint	−97.86	1853.00	46.77
Diff Mid-to-L Wrist	−583.91	69.00	327.31
Diff Mid-to-R Wrist	544.29	275.00	303.18

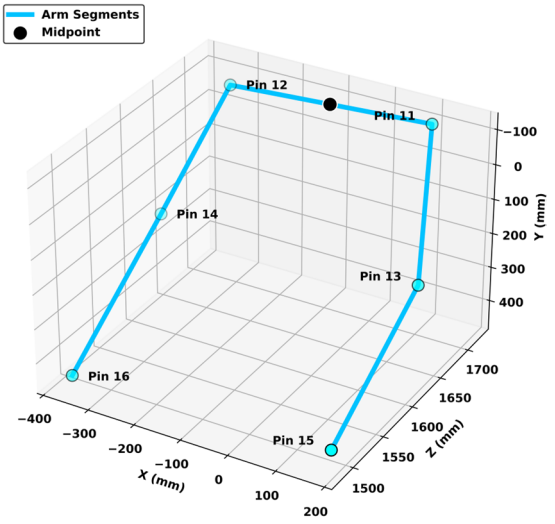


Fig. 12. Midpoint & wrist vectors (a) where Pin 15 (L wrist) and Pin 16 (R wrist) are colored red.

TABLE V. RAW LANDMARKS COORDINATE

Landmark	X	Y	Z
11 (L shoulder)	99.19	-99.23	1721.00
13 (L Elbow)	174.61	211.91	1639.00
15 (L Wrist)	175.06	446.97	1496.00
12 (R shoulder)	−305.09	−106.68	1696.00
14 (R Elbow)	−373.16	202.33	1631.00
16 (R Wrist)	−360.02	440.49	1483.00

Fig. 13 presents the result of the coordinate transformation. A stable reference point—the shoulder midpoint—was calculated and located at (dX: −102.95, dY: −102.95, dZ: 1708.50) as shown in Table VI. The 3D plot then depicts the vectors extending from this midpoint to both wrists, with corresponding unscaled vector data: left wrist (dX: −278.01, dY: −549.92, dZ: 212.50) and right wrist (dX: 257.07, dY: −543.45, dZ: 225.50). This transformation effectively expressed the hand positions within a local, torso-centered coordinate system, isolating arm motion for robotic control.

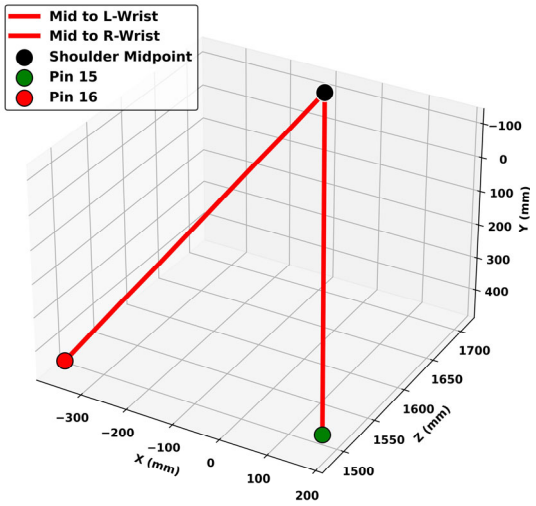


Fig. 13. Midpoint & wrist vectors (b) where Pin 15 (L wrist) and Pin 16 (R wrist) are colored green and red, respectively.

TABLE VI. UNSCALED VECTOR DATA

Measurement	dX	dY	dZ
Midpoint	−102.95	−102.95	1708.50
L Wrist Vector	−278.01	−549.92	212.50
R Wrist Vector	257.07	−543.45	225.50

Through this process, human hand movements were proportionally reproduced by the robot while ensuring compliance with mechanical safety limits.

As illustrated in Fig. 14, the scaling process was applied along all three motion axes. Along the Y-axis (vertical), human hand movements from 0 to 600 mm were proportionally reduced to the robot's range of 0–473 mm (Fig. 14(a)). For the X-axis (lateral), motions within 350–740 mm were mapped to 187–566 mm (Fig. 14(b)). Along the Z-axis (depth), movements spanning 0–580 mm were scaled to 0 to –473 mm (Fig. 14(c)). In every case, safety boundaries were imposed so that movements beyond the tracking range caused position saturation at the robot's mechanical limits, thereby ensuring safe operation.

The final stage of the motion-processing pipeline produced scaled target vectors for both arms, which were transmitted to the robot's Inverse Kinematics (IK) solver. A 3D plot visualized the resulting hand trajectories within the robot's coordinate frame, with the scaled left-hand path (green) and right-hand path (red) representing the filtered and scaled movements of the human operator (Fig. 15). The corresponding numerical data were recorded, showing final scaled target coordinates of (X: −187.0, Y: 125.34,

Z: −447.05) for the left arm and (X: 187.0, Y: 135.59, Z: −441.77) for the right arm (Table VII). These results confirmed the successful conversion of dynamic human motion into accurate, safe, and executable 3D targets for humanoid robotic control.

After detecting and scaling the user's pose landmarks to the robot's workspace, inverse kinematics was applied to convert the 3D coordinates into corresponding servo motor angles for the dual-arm robot. System performance was continuously monitored through a real-time dashboard that displayed diagnostic parameters such as arm status (idle or moving), servo velocity, step value, and the difference between target and current joint angles (Fig. 16).

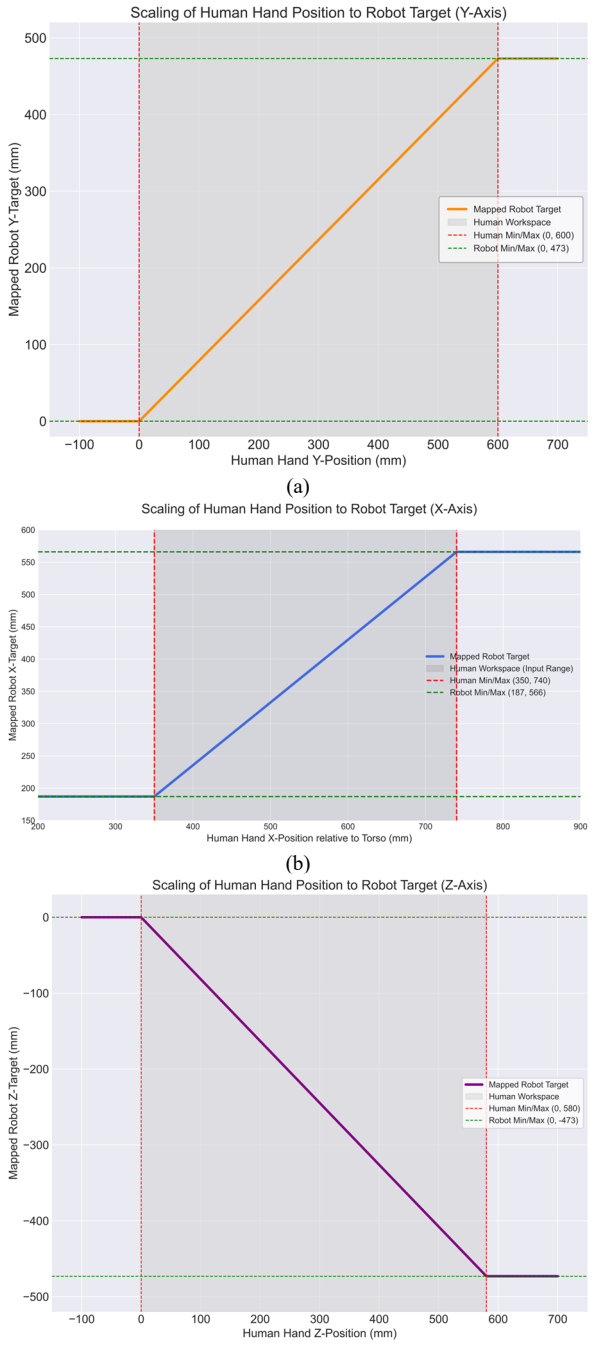


Fig. 14. Scaling of human hand positions relative to robot target in (a) Y; (b) X; and (c) Z axes respectively.

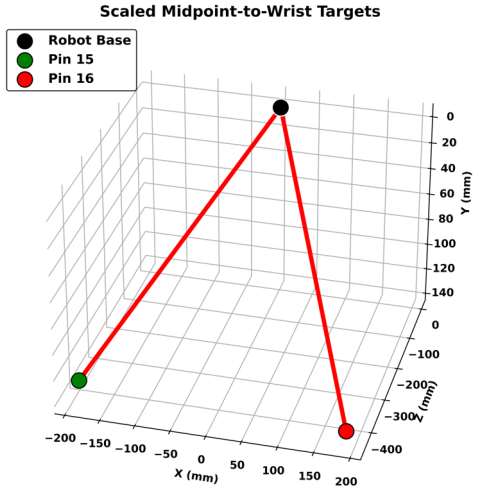


Fig. 15. Scaled targets for inverse kinematics.

TABLE VII. FINAL SCALED COORDINATES

Arm	Scaled X	Scaled Y	Scaled Z
Target Pin 15 (Left)	−187.00	125.34	−447.05
Target Pin 16 (Right)	187.00	135.59	−441.77

The captured results demonstrate the system's capability to monitor both arms independently. In the first instance, the right arm was idle (green status) with a servo velocity of 0.0 mm/s (step 1) and complete alignment between target and actual joint angles, while the left arm was active (red status) with a velocity of 89.1 mm/s (step 8) and visible deviation between target and current angles (Fig. 16). This independent, real-time tracking of each arm's velocity and step value was essential for validating accurate and asynchronous motion execution.

The complete kinematic methodology of the dual-arm robotic system was calculated to define its structure, target acquisition, and corresponding mathematical formulation. The kinematic control process began with the robot's perception of its environment, where the RGB-D camera captured target coordinates—shown as magenta and cyan stars in Fig. 17—that served as the desired end-effector positions. These targets were listed in Table VIII, specifying for instance the left arm's goal at (X: −209.050, Y: 249.229, Z: −3.087) and the right arm's at (X: 336.886, Y: 1.538, Z: −88.104).

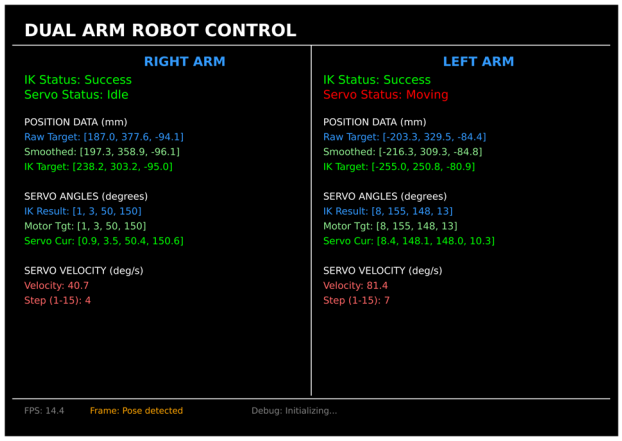


Fig. 16. System dashboard.

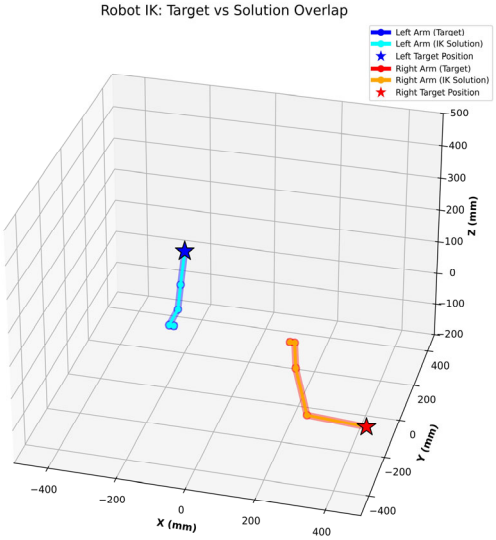


Fig. 17. Solutions of the kinematic model.

To reach these targets, the system utilized a predefined geometric model defined by the DH parameters (Table IX), which described each joint through its link length (a), offset (d), and twist angle (α). For example, Joint 0 of the left arm was defined by a link length of −173.4 mm and a twist of −57°. This model provided the structural basis for the IK solver to compute feasible joint angles within mechanical limits.

Once the geometric model and targets were defined, the IK solver computed the required joint configuration, visualized in Fig. 17 as the solved pose (IK). The resulting 4×4 homogeneous transformation matrix, presented in Table X, numerically verified the end-effector's position and orientation.

The "Position" column of this matrix corresponded precisely to the initial target coordinates, confirming that the IK computation successfully translated perception into an accurate and achievable robotic motion.

TABLE VIII. TARGET 3D COORDINATES FOR LEFT AND RIGHT ARM COORDINATION

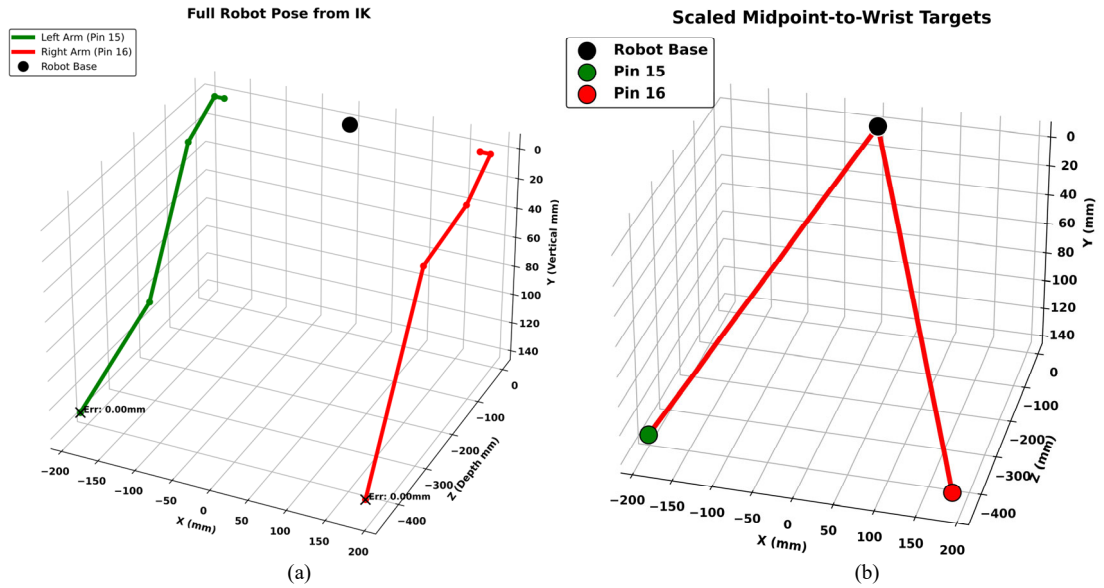
Coordinate	Left Arm Target Position	Right Arm Target Position
X	−209.050	336.886
Y	249.229	1.538
Z	−3.087	−88.104

TABLE IX. GEOMETRIC AND KINEMATIC PARAMETERS FOR THE LEFT AND RIGHT ROBOT ARMS BASED ON THE DH CONVENTION

Joint	Left Arm DH Parameters				Right Arm DH Parameters			
	a (mm)	d (mm)	alpha (deg.)	Limits	a (mm)	d (mm)	alpha (deg.)	Limits
Joint 0	−173.4	0	−57	0°–180°	173.4	0	57	0°–180°
Joint 1	−13.7	0	57	70°–180°	13.7	0	57	0°–96°
Joint 2	−91.4	0	0	0°–90°	91.4	0	0	0°–180°
Joint3	−159.4	0	0	0°–90°	159.4	0	0	0°–180°

TABLE X. 4×4 HOMOGENEOUS TRANSFORMATION MATRICES FROM THE INVERSE KINEMATICS SOLVER

ARM IK Results	X	Y	Z	Position
Left Arm IK Results	0.881	0.45	−0.13	−209.05
	−0.457	0.89	0	249.22
	0.124	0.06	0.99	−3.08
	0.000	0	0	1000
Right Arm IK Results	0.21	−0.97	−0.001	336.88
	−0.01	−0.003	−1000	153
	0.97	0.21	−0.01	−88.10
	0	0	0	1000



(a) (b)

Fig. 18. IK validation (a) IK pose; (b) torso-anchored targets.

TABLE XI. PERFORMANCE EVALUATION OF THE PROPOSED IK TECHNIQUE

Metric	Left Arm	Right Arm
Trials	1	1
Best Trial	0	0
Success	TRUE	TRUE
Error	0.5	0.5

To validate the IK solver, its precision in converting scaled target coordinates into accurate robotic poses was analyzed. The 3D plots in Fig. 18(a) and (b) illustrate this transformation, showing how human motion inputs were reconstructed into feasible robotic trajectories. On the left, the full robot possessing base, left, and right arm pathways—was reconstructed in millimeter-scale space, confirming symmetry and ergonomic feasibility. On the right, the scaled midpoint-to-wrist vectors were plotted relative to the torso base, with Pins 15 and 16 denoting the left and right wrists. These visualizations demonstrate the normalization of human pose data into stable, robot-centric control targets for dual-arm replication. Quantitative validation further confirmed accuracy: Table XI reported a "TRUE" success status with minimal error margins (0.5), and Table XII showed near-perfect agreement between target and achieved positions. For example, the left arm target (X: −187.00, Y: 125.34, Z: −447.05) was reproduced precisely. Collectively, these results verify the IK solver's high reliability and precision in robotic pose replication.

TABLE XII. TARGET VERSUS. ACHIEVED POSITION COMPARISON

Position Type	X	Y	Z
Left Target	−187.00	125.34	−447.05
Left Result	−187.00	125.34	−447.55
Right Target	187.00	136.5	−441.77
Right Result	187.00	136	−441.77

TABLE XIII. FINAL JOINTS CONFIGURATION ANGLES

Joint	Left Arm (degrees)	Right Arm (degrees)
Joint 0	0	0.09
Joint 1	92.1	83.28
Joint 2	115.95	49.2
Joint 3	0	0

The data in Table XIII presents the final joint configuration angles for both arms highlights the precision of the inverse kinematics algorithm in computing joint movements required to reach target positions. The calculated angles—0°, 92.1°, 115.95°, and 0° for the left arm, and 0.09°, 83.28°, 49.2°, and 0° for the right arm—demonstrate accurate correspondence between the commanded and achieved positions. This four-Degree-of-Freedom (DOF) configuration confirms the algorithm's effectiveness in positioning and orienting the robotic limbs with high accuracy.

Fig. 19 illustrate the system's ability to translate continuous human arm motion into discrete speed commands for robotic control. The subject raw wrist velocities shown in Fig. 19(a) (left arm) and Fig. 19(b) (right arm) indicate non-uniform movement characterized by rapid bursts, with peaks exceeding 600 mm/s and 550 mm/s, respectively. This continuous velocity data was converted into quantized speed steps ranging from 1 to 15, which directly defined the robot's motion commands (Fig. 19(c) and (d)). A clear similarity was observed—high-velocity peaks in Fig. 19(a) aligned with maximum speed commands (step 15) in Fig. 19(c), while stationary periods at 0 mm/s produced a constant step 1, indicating idle motion. These results demonstrate a robust and responsive mapping process that effectively interprets human movement dynamics into structured, executable robotic control signals.

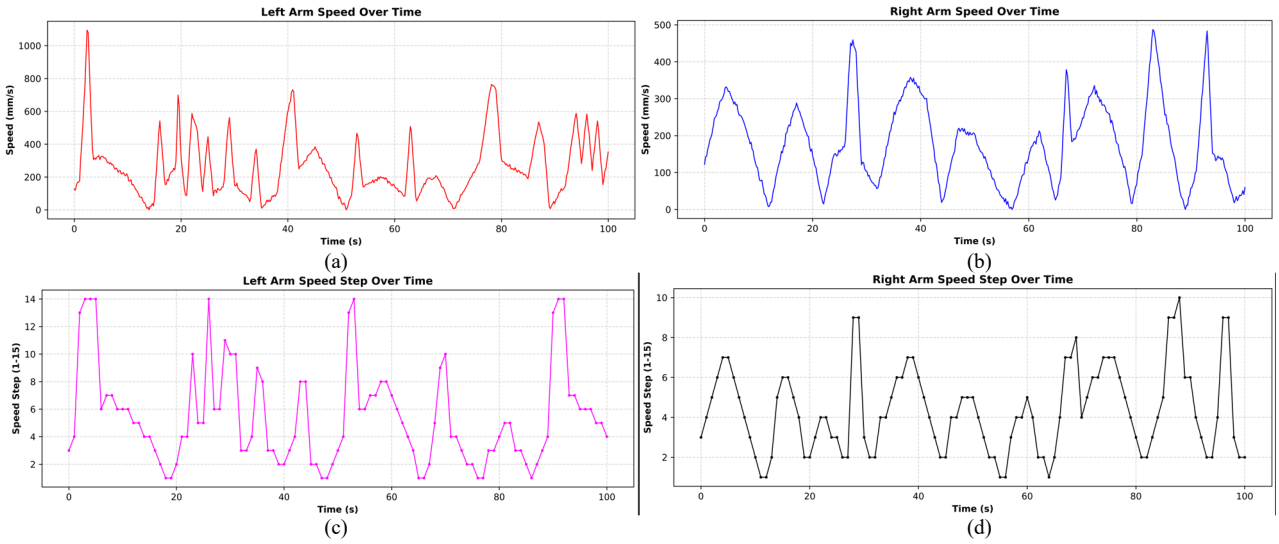


Fig. 19. Human–robot motion and speed response analysis. (a) left wrist motion; (b) right wrist motion; (c) left end-effector speed; (d) right end-effector speed.

Experimental tests demonstrated the system's capability to accurately replicate human arm movements within predefined spatial limits (Fig. 20). Images a to c show single-arm motion, while d to f images illustrate dual-arm operation. These figures depict a user interacting with the fixed camera and the ability to emulate natural human gestures in a wide range of motion. The design was optimized for ergonomic performance, ensuring suitability

for daily activities. Movement stabilization algorithms played a key role in minimizing jitter and enhancing responsiveness—critical for real-time applications. This stability is reflected in Fig. 20, where coordinate responses along the X, Y, and Z axes are presented in corresponding pairs (Fig. 20(a) and (d), Fig. 20(b) and e, Fig. 20(c) and (f)).

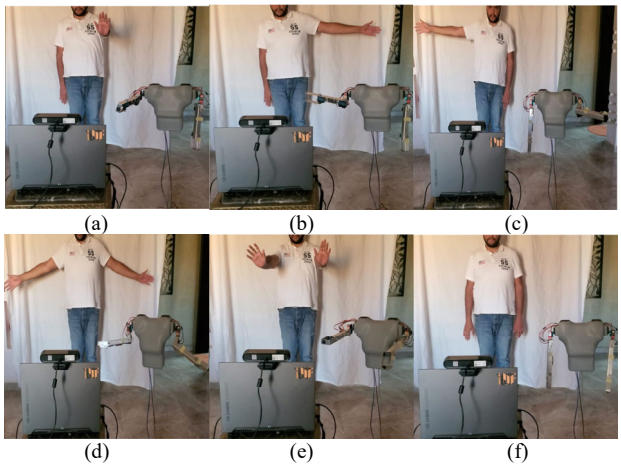


Fig. 20. Experimental testing of the pose estimation mimicry system. (a) right-arm pose in Z-axis; (b) right-arm pose in X-axis; (c) left-arm pose in X-axis; (d) two-arm pose in X-axis; (e) two-arm pose in Z-axis; (f) two-arm pose in Y-axis.

IV. Discussion

The developed humanoid mimicry system successfully integrated hardware, motion capture, signal transformation, and robotic control into a unified framework. The 3D-printed torso, dual robotic arms, control laptop, and RGB-D camera functioned reliably, with mechanical alignment and motion verified through preliminary testing [36]. These findings align with previous studies emphasizing precision assembly to reduce control errors and improve reliability [37]. Continuous monitoring through a real-time dashboard tracked joint angles, positions, and servo velocities, ensuring transparency and rapid deviation detection—consistent with closed-loop feedback principles in robotics [38]. The near-instantaneous correspondence between human and robotic motion confirmed the stability of the complete pipeline, spanning marker less tracking, torso-centers transformation, scaling, and inverse kinematics [1, 36].

Real-time landmark extraction validated accurate skeletal reconstruction using MediaPipe, which detected key anatomical points (shoulders, elbows, wrists) from RGB-D input [32]. The resulting millimeter-scale coordinates established a quantitative base for control [39]. Derived biomechanical features—such as the shoulder midpoint and wrist vectors—further enhanced stability and interpretability, generating structured datasets suited for robotic execution [40, 41].

A two-stage transformation converted human poses into a torso-anchored coordinate frame. The shoulder midpoint provided a stable reference for computing wrist vectors, isolating arm motion relative to the torso, and improving robustness against camera placement. This approach agrees with established motion-capture and robotics practices emphasizing body-centric reference frames for stability and consistent inverse kinematics mapping [42–44].

Following filtering and scaling, the system generated bound arm trajectories constrained within the robot's workspace, ensuring safe operation and symmetrical motion. This normalization process was crucial for producing refined inputs to the IK module and avoiding joint-limit violations [45–47]. Real-time monitoring demonstrated the system's ability to execute asynchronous dual-arm commands with high precision. Closed-loop feedback allowed rapid detection and correction of angular deviations, maintaining safe and transparent performance consistent with established teleoperation standards [48–50]. The IK solver consistently reported "TRUE" success, with an average positioning error of only 0.5, resulting in accurate and collision-free joint configurations. The close correspondence between commanded and actual positions confirmed its robustness and reliability for dual-arm teleoperation and rehabilitation applications [50–52].

Recent tests also demonstrated smooth reproduction of natural reaching motions, supported by stabilization algorithms that mitigated jitter and occlusion effects. These results, consistent with prior work on adaptive control and kinematic stability [53], highlight the scalability and potential of the system for assistive and prosthetic robotics, where intuitive and stable control is critical [54].

V. Conclusion

This work introduced a real-time humanoid mimicry framework that maps human upper-limb motion to dual robotic arms using MediaPipe Pose and an RGB-D camera. The integrated pipeline—combining pose detection, coordinates transformation, workspace scaling, and inverse kinematics—achieved precise and responsive replication of human movements. Experimental validation confirmed high accuracy, with an average wrist-tracking RMSE of 12.4 mm, a latency of 86 ms, and minimal inverse kinematics error (≤ 0.5 mm). The system demonstrated safe, synchronized dual-arm operation through real-time dashboard monitoring.

The proposed setup provides a low-cost, marker-less alternative to traditional motion-capture systems, suitable for education, rehabilitation, and teleoperation. The system's limitation occurs at greater distances from the camera, where the human subject becomes pixelated in the video feed, preventing MediaPipe from recognizing the human or detecting key points accurately. Future work will focus on multi-camera integration for improved depth estimation, adaptive control through learning-based methods, and enhanced hardware with torque feedback to enable complex manipulation and physical interaction. Expanding the framework to full-body tracking and immersive telepresence is also planned.

Conflict of Interest

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

All authors contributed to the conception and design of the study. Material preparation, data collection, and analysis were carried out by OS, AA, and MRAA. OS drafted the initial manuscript, and all authors provided feedback on earlier versions; all authors had approved the final version.

REFERENCES

[1] M. Raslan, A. Abdellatif and M. R. A. Atia, "A novel exoskeleton finger mechanism for robotic applications," in *Proc. 2021 3rd Novel Intelligent and Leading Emerging Sciences Conf. (NILES)*, Giza, 2021, pp. 5–10. doi: 10.1109/NILES53778.2021.9600087

[2] A. Ibrahim, A. Ammounah, S. Alfayad *et al.,* "Hydraulic robotic leg for HYDROïD robot: Modeling and control," *J. Robot. Mechatron.*, vol. 34 no. 3, pp. 576–587, 2022.

[3] U. Othman and E. Yang, "Human–robot collaborations in smart manufacturing environments: Review and outlook," *Sensors*, vol. 23, no. 12, 5663, 2023. doi: 10.3390/s23125663

[4] A. Halim, A. A. Ibrahim, M. I. Awad *et al.,* "Optimization of sensor number for lower limb prosthetics using genetic algorithm," in *Proc. 2020 International Conf. on Innovative Trends in Communication and Computer Engineering (ITCE)*, Aswan, 2020, pp. 210–215. doi: 10.1109/ITCE48509.2020.9047780

[5] M. Abdelbar, I. Mohamed, A. Abdellatif *et al.,* "Towards the mechatronic development of a new upper-limb exoskeleton (SAMA)," *Designs*, vol. 6, no. 5, 80, 2022. https://doi.org/10.3390/designs6050080

[6] J. Kim, J.-H. Ahn, and Y. Kim, "Immersive interaction for inclusive virtual reality navigation: Enhancing accessibility for socially underprivileged users," *Electronics*, vol. 14, no. 5, 1046, 2025. doi: 10.3390/electronics14051046

[7] R. Damaševičius and T. Sidekerskienė, "Virtual worlds for learning in metaverse: A narrative review," *Sustainability*, vol. 16, no. 5, 2032, 2024. doi: 10.3390/su16052032

[8] A. Strzelecki, B. Kolny, and M. Kucia, "Smart homes as catalysts for sustainable consumption: A digital economy perspective," *Sustainability*, vol. 16, no. 11, 4676, 2024. doi: 10.3390/su16114676

[9] C. Dong and G. Du, "An enhanced real-time human pose estimation method based on modified YOLOv8 framework," *Scientific Reports*, vol. 14, 8012, 2024. doi: 10.1038/s41598-024-58146-z

[10] E. Samkari, M. Arif, M. Alghamdi *et al.,* "Human pose estimation using deep learning: A systematic literature review," *Machine Learning and Knowledge Extraction*, vol. 5, pp. 1612–1659, 2023.

[11] R. Sun, Z. Lin, S. Leng *et al.,* "An in-depth analysis of 2D and 3D pose estimation techniques in deep learning: Methodologies and advances," *Electronics*, vol. 14, no. 7, 1307, 2025. doi: 10.3390/electronics14071307

[12] S. Dill, A. Ahmadi, M. Grimmer *et al.,* "Accuracy evaluation of 3D pose reconstruction algorithms through stereo camera information fusion for physical exercises with MediaPipe Pose," *Sensors*, vol. 24, no. 23, 7772, 2024. doi: 10.3390/s24237772

[13] Z. Hu, C. Zhang, X. Wang *et al.,* "Light-adaptive human body key point detection algorithm based on multi-source information fusion," *Sensors (Basel)*, vol. 24, 2024. doi: 10.3390/s24103021

[14] Y. Jang, I. Jeong, M. Y. Heravi *et al.,* "Multi-camera-based human activity recognition for human–robot collaboration in construction," *Sensors*, vol. 23, 6997, 2023.

[15] H. Said, K. Mahar, S. E. Sorour *et al.,* "IMITASD: Imitation assessment model for children with autism based on human pose estimation," *Mathematics*, vol. 12, 3438, 2024.

[16] R. Onciul, C.-I. Tataru, A. V. Dumitru *et al.,* "Artificial intelligence and neuroscience: Transformative synergies in brain research and clinical applications," *Journal of Clinical Medicine*, vol. 14, no. 2, 550, 2025. doi: 10.3390/jcm14020550

[17] A. Garmendia-Orbegozo, M. A. Anton, and J. D. Nuñez-Gonzalez, "Reduction of vision-based models for fall detection," *Sensors*, vol. 24, 7256, 2024.

[18] J. Gutiérrez, V. Rodríguez, and S. Martin, "Comprehensive review of vision-based fall detection systems," *Sensors*, vol. 21, 947, 2021.

[19] W.-C. Lin, Y.-C. Tu, H.-Y. Lin *et al.,* "A comparison of deep learning techniques for pose recognition in up-and-go pole walking exercises using skeleton images and feature data," *Electronics*, vol. 14, 1075, 2025.

[20] J.-W. Kim, J.-Y. Choi, E.-J. Ha *et al.,* "Human pose estimation using MediaPipe pose and optimization method based on a humanoid model," *Applied Sciences*, vol. 13, 2700, 2023.

[21] F. Roggio, B. Trovato, M. Sortino *et al.,* "A comprehensive analysis of the machine learning pose estimation models used in human movement and posture analyses: A narrative review," *Heliyon*, vol. 10, e39977, 2024. doi: 10.1016/j.heliyon. 2024.e39977

[22] M. Desai and H. Mewada, "A novel approach for yoga pose estimation based on in-depth analysis of human body joint detection accuracy," *PeerJ Computer Science*, vol. 9, e1152, 2023. doi: 10.7717/peerj-cs.1152

[23] A. Upadhyay, N. K. Basha, and B. Ananthakrishnan, "Deep learning-based yoga posture recognition using the Y_PN-MSSD model for yoga practitioners," *Healthcare*, vol. 11, 609, 2023.

[24] Y. Zhou, F. N. Rashid, M. M. Daud *et al.,* "Machine learning-based computer vision for depth camera-based physiotherapy movement assessment: A systematic review," *Sensors (Basel)*, vol. 25, no. 5, 1586, 2025. doi: 10.3390/s25051586

[25] A. Akbari, F. Haghverd, and S. Behbahani, "Robotic home-based rehabilitation systems design: From a literature review to a conceptual framework for community-based remote therapy during COVID-19 pandemic," *Frontiers in Robotics and AI*, vol. 8, 2021. doi: 10.3389/frobt.2021.612331

[26] B. Cunha, R. Ferreira, and A. S. P. Sousa, "Home-based rehabilitation of the shoulder using auxiliary systems and artificial intelligence: An overview," *Sensors*, vol. 23, 7100, 2023.

[27] C. S. Stafie, I. G. Sufaru, C. M. Ghiciuc *et al.,* "Exploring the intersection of artificial intelligence and clinical healthcare: A multidisciplinary review," *Diagnostics (Basel)*, vol. 13, 2023. doi: 10.3390/diagnostics13121995

[28] Y. Guo, T. Gao, A. Dong *et al.,* "A survey of the state of the art in monocular 3D human pose estimation: Methods, benchmarks, and challenges," *Sensors*, vol. 25, 2409, 2025.

[29] J.-Y. Choi, E. Ha, M. Son *et al.,* "Human joint angle estimation using deep learning-based three-dimensional human pose estimation for application in a real environment," *Sensors*, vol. 24, 3823, 2024.

[30] S. Ding and F. Zhu, "Distributed robust filtering over sensor networks with quantized measurement and switching topologies," *Electronics*, vol. 10, 2336, 2021.

[31] J. P. Angelo, S. J. Chen, M. Ochoa *et al.,* "Review of structured light in diffuse optical imaging," *Journal of Biomedical Optics*, vol. 24, pp. 1–20, 2018. doi: 10.1117/1.JBO.24.7.071602

[32] S. Tang, Q. Zhu, W. Chen *et al.,* "Enhanced RGB-D mapping method for detailed 3D indoor and outdoor modeling," *Sensors*, vol. 16, 1589, 2016.

[33] C. Wongoutong, "The impact of neglecting feature scaling in k-means clustering," *PLoS One*, vol. 19, e0310839, 2024. doi: 10.1371/journal.pone.0310839

[34] H. Yin, H. Jin, Y. Peng *et al.,* "Design of a humanoid upper-body robot and trajectory tracking control via ZNN with a matrix derivative observer," *Biomimetics (Basel)*, vol. 10, 2025. doi: 10.3390/biomimetics10080505

[35] G. C. Daraviña, J. L. Valencia, G. A. Holguin *et al.,* "Visual servoing and Kalman filter applied to parallel manipulator 3-RRR," *Electronics*, vol. 13, 2703, 2024.

[36] S. Garcia, F. Gomez-Donoso, and M. Cazorla, "Enhancing human–robot interaction: Development of multimodal robotic assistant for user emotion recognition," *Applied Sciences*, vol. 14, 11914, 2024.

[37] R. Liu, V. I. Sucala, M. Luis *et al.,* "Systematic review of service quality models in construction," *Buildings*, vol. 15, 2331, 2025.

[38] Q. Xiang, C. Chen, and Y. Jiang, "Servo collision detection control system based on robot dynamics," *Sensors*, vol. 25, 1131, 2025.

[39] Y. Wang, B. Hussain, and C. P. Yue, "VLP landmark and SLAM-assisted automatic map calibration for robot navigation with semantic information," *Robotics*, vol. 11, 84, 2022. doi: 10.3390/robotics11040084

[40] Y. Dong and S. Payandeh, "Hand kinematic model construction based on tracking landmarks," *Applied Sciences*, vol. 15, 8921, 2025. https://doi.org/10.3390/app15168921

[41] A. Nussibaliyeva, G. Sergazin, G. Tursunbayeva *et al.,* "Development of an artificial vision for a parallel manipulator using machine-to-machine technologies," *Sensors*, vol. 24, 3792, 2024. https://doi.org/10.3390/s24123792

[42] R. Maskeliūnas, R. Damaševičius, T. Blažauskas *et al.,* "BiomacVR: A virtual reality-based system for precise human posture and motion analysis in rehabilitation exercises using depth sensors," *Electronics*, vol. 12, 339, 2023. https://doi.org/10.3390/electronics12020339
[43] S.-H. Tseng, J.-C. Chiang, C.-E. Shiue *et al.,* "Integrating OpenPose for proactive human–robot interaction through upper-body pose recognition," *Electronics*, vol. 14, 3112, 2025. https://doi.org/10.3390/electronics14153112
[44] Q. Li, F. Cicirelli, A. Vinci *et al.,* "Quadruped robots: Bridging mechanical design, control, and applications," *Robotics*, vol. 14, 57, 2025. https://doi.org/10.3390/robotics14050057
[45] Y. Zheng, H. Zhang, G. Zheng *et al.,* "Research on motion transfer method from human arm to bionic robot arm based on PSO-RF algorithm," *Biomimetics*, vol. 10, no. 6, 392, 2025. https://doi.org/10.3390/biomimetics10060392
[46] I. F. Astorquia, G. Villate-Castillo, A. Tellaeche *et al.,* "Comparative benchmark of sampling-based and DRL motion planning methods for industrial robotic arms," *Sensors*, vol. 25, 5282, 2025. https://doi.org/10.3390/s25175282
[47] Q. Zhang, Q. Liu, J. Duan *et al.,* "Research on teleoperated virtual reality human–robot five-dimensional collaboration system," *Biomimetics (Basel)*, vol. 8, 2023. doi: 10.3390/biomimetics8080605
[48] C. Li, Z. Liu, L. Li *et al.,* "Improved PPO optimization for robotic arm grasping trajectory planning and real-robot migration," *Sensors*, vol. 25, 5253, 2025. https://doi.org/10.3390/s25175253
[49] X. Jiang, B. Wu, S. Li *et al.,* "Multi-humanoid robot arm motion imitation and collaboration based on improved retargeting," *Biomimetics (Basel)*, vol. 10, 2025. doi: 10.3390/biomimetics10030190
[50] Y. Zhu, T. Wang, C. Wang *et al.,* "Complexity-driven trust dynamics in human–robot interactions: Insights from AI-enhanced collaborative engagements," *Applied Sciences*, vol. 13, 12989, 2023. https://doi.org/10.3390/app132412989
[51] M. Zhang, B. Qian, J. Gao *et al.,* "Recent advances in portable dry electrode EEG: Architecture and applications in brain–computer interfaces," *Sensors*, vol. 25, 5215, 2025. https://doi.org/10.3390/s25165215
[52] R. Li and L. Xu, "Active RIS-Assisted uplink NOMA with MADDPG for remote state estimation in wireless sensor networks," *Sensors*, vol. 25, 4878, 2025. doi: 10.3390/s25154878
[53] P. Jha, G. P. K. Yadav, and D. Bandhu, "Human–machine interaction and implementation on the upper extremities of a humanoid robot," *Discover Applied Sciences*, vol. 6, 152, 2024. doi: 10.1007/s42452-024-05734-3
[54] H. M. K. K. M. B. Herath, N. Madusanka, S. L. P. Yasakethu *et al.,* "Biomimetic robotics and sensing for healthcare applications and rehabilitation: A systematic review," *Biomimetics*, vol. 10, no. 7, 466, 2025. doi: 10.3390/biomimetics10070466