Telepresence Robotic Arm for Remote Dermatological Monitoring of Infectious Wards

Tanay Raghunandan Srinivasa
Robotics and Autonomous Systems
Plaksha University
Mohali, India
tanay.srinivasa@plaksha.edu.in

Jia Bhargava
Robotics and Autonomous Systems
Plaksha University
Mohali, India
jia.bhargava@plaksha.edu.in

Moksh Soni
Robotics and Autonomous Systems
Plaksha University
Mohali, India
moksh.soni@plaksha.edu.in

***Abstract*—In this paper, we design and integrate a telepresence robotic arm system specifically for remote dermatological monitoring within infectious wards. This development is motivated by the critical need to reduce the risk of nosocomial infections for both susceptible patients and medical practitioners. The proposed system utilizes a Temi 3 robot equipped with a custom 2-degree-of-freedom robotic arm mounted on the Temi tray. We present the specific design considerations, requirements, and mechanical optimizations of the robotic arm. Additionally, we implement and compare various computer vision techniques for patient identification, object tracking, face tracking, and skin anomaly detection. These techniques are benchmarked by measuring and comparing their inference latency to validate the feasibility of real-time deployment. Finally, we present considerations and discussions regarding future iterations of this system to enhance its efficacy in clinical environments. Our open-source code can be found at: https://github.com/tanayrs/telepresence-medical-bot**

***Index Terms*—Telepresence Robots, Medical Robotics, Robotic Arm**

I. Introduction

Hospital-acquired infections (HAIs), also known as nosocomial infections, represent a critical public health challenge, imposing a severe burden on healthcare systems globally. According to recent reports from the World Health Organization (WHO), HAIs affect hundreds of millions of patients annually, with a significantly higher prevalence in low- and middle-income countries (LMICs) compared to high-income nations. Data indicates that approximately 7% of hospitalized patients in developed countries and up to 15% in developing countries acquire at least one HAI during their stay [1]. The mortality associated with these infections is alarmingly high; in Europe alone, HAIs are responsible for approximately 37,000 attributable deaths annually, while in the United States, they account for nearly 99,000 deaths per year [2]. In resource-limited settings such as sub-Saharan Africa, the impact is even more devastating, with infection-associated mortality rates reaching as high as 22% [3]. These figures underscore the urgent need for interventions that can sever the chain of transmission within clinical environments.

This project has been funded by the Robotics Courses Fall 2025 Fund under Plaksha University. We would like to thank Prof. Sunita Chauhan, Prof. Sandeep Manjanna and Deepanshu Sharma for their assistance during the process of this project.

A primary cause for the spread of nosocomial pathogens is the necessary physical proximity between medical practitioners and patients. Healthcare workers (HCWs) are not only at an elevated risk of contracting infections but also serve as potential carriers, transmitting pathogens between highly susceptible immunocompromised patients. During the COVID-19 pandemic, this vulnerability was starkly highlighted; thousands of frontline workers were infected. Studies in regions like Egypt show HCW infection rates as high as 3.1% even prior to the pandemic peak, with physicians often facing the highest relative risk due to frequent direct contact [4]. The transmission of multidrug-resistant organisms (MDROs) via direct contact remains a persistent threat, with contact transmission being the most common route for pathogens such as MRSA and C. *difficile* [5]. Consequently, reducing physical exposure without compromising the quality of care has become a priority in infectious disease management.

To address this gap, evaluates the feasibility and details the development of a telepresence robot integrated with a robotic arm specifically designed for dermatological monitoring in infectious wards. Dermatology relies heavily on visual inspection, which traditionally requires close physical proximity, thereby increasing the risk of pathogen exposure. By utilizing a telerobotic system, medical practitioners can conduct thorough visual examinations and manipulate inspection tools remotely, effectively eliminating the need for direct physical contact. This approach aims to safeguard healthcare personnel from becoming victims of HAIs while simultaneously protecting vulnerable patients from cross-contamination, ultimately reducing the incidence of nosocomial transmission in high-risk infectious wards.

The rest of this paper is organized as follows: In Section II we discuss the design requirements and considerations of our monitoring system, in Section III we discuss our results, in Section IV we discuss limitations and future work, and we conclude this paper in Section V.

II. Methodology

In this section, we discuss the design requirements, considerations, and interface of the monitoring system.

A. Temi

We choose the Temi 3 telepresence robot as the base for our robotic arm. Temi comes in-box with the ROBOX^{TM} navigation system, that is their propriety human-robot interaction and autonomous navigation system. Additionally, Temi comes with in-built 2-D mapping, localization, obstacle avoidance and path planning capabilities. We levarage these capabilities to navigate autonomously between hospital rooms.

The Temi robot gives us access to navigate between locations through their Java SDK, and we use the 'pytemi' public library to interact with this interface. The functionalities of the Temi robot can be seen in Table I.

TABLE I
TEMI ROBOT FUNCTIONALITIES

Function Name	Functionality
rotate()	Rotate about the systems base.
tilt()	Tilt the display of the Temi robot.
follow()	Enable Temi to follow a target.
goto()	Navigate to a particular location.
stop()	Emergency stop during Temi's navigation.
joystick()	Fine control of x-y motion of Temi.
tts()	Text-to-Speech.
video()	Play a video at a particular URL
webview()	Open a particular URL

Due to the closed nature of the Temi ecosystem, we are unable to access any of Temi's onboard sensors directly. Additionally, Temi's movement between locations cannot be altered with a custom social navigation policy. During preliminary tests, we noticed that Temi's navigation policy easily gets confused and moves in circles if interrupted by moving obstacles.

Given these capabilities and constraints, we use Temi's goto() function to move between hospital rooms and key locations, and use the joystick() control for fine placement of the robot next to the patient's bed. After this phase, we use the robotic arm for inspection.

B. Robotic Arm

For the task of skin inspection, we need a 3 degree-of-freedom (DoF) robotic arm. With Temi's rotate() function and a 2 DoF arm mounted on Temi about its rotational axis, we can achieve an effective 3 DoF arm. Given this simplification, the basic configuration of the arm can be seen in Fig. 1. The payload for this robotic arm is a high-resolution camera mounted at the end of the arm (point C). We choose the Frontech Webcam 2255 for the arm as it is lightweight (100g) and captures a high-definition (720p) video feed.

The primary design consideration has to do with the torques encountered by the motors driving the arms. To minimize this torque, we choose to mount both arm motors at the base of the arm (point A) to prevent the weight of the top arm motor from increasing the torque load on the base motor. This configuration contrasts with traditional designs where motors are mounted at both the base (point A) and the elbow joint (point B).

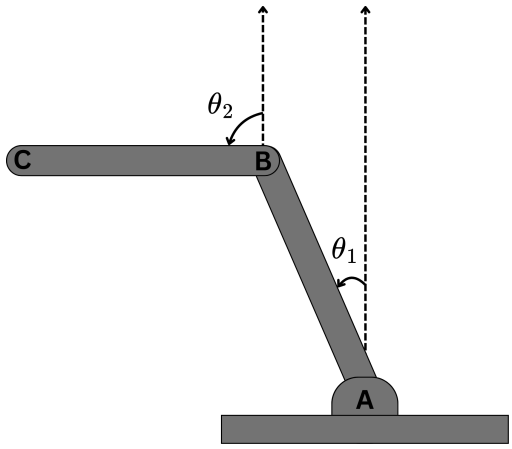

Fig. 1. Basic Configuration of the 2-DoF Robotic Arm

The next design consideration comes from the size of the patients bed. In this study, we take the size of the bed to be 4 feet. Consequently, the arm requires a maximum reach of 2 feet. We also limit the maximum angle of the base arm to 20 degrees from the normal. For convenience, we set both arms to have equal length, arriving at the total length of each arm as 350mm.

We choose the NEMA-17 stepper motor rated at 6.5 kg·cm torque as the motors for both arms. Given the torque calculations presented in the Appendix3, we add a 6:1 gear reduction for the base arm, and a 3:1 gear reduction for the top arm. The top arm and base arm motors are operated at 1.6 A and 2 A respectively with 1/8th stepping through TB6560 stepper motor drivers.

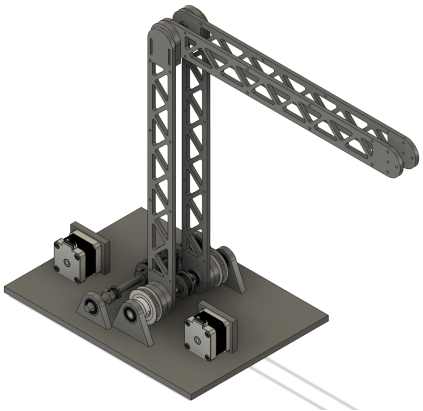

Fig. 2. Computer Aided Model of Fully Assembled Robot Arm

The final arm design can be seen in Fig. 2. The truss-like design was cut into the arms to reduce the weight, while maintaining structural rigidity. The motors were shifted off axis to allow the decoupling of the base arm and upper axes, while maintaining independent control. This arrangement of motors also allows the input of absolute angles as the arm configurations, that is, the upper arm automatically maintains its angle θ_2 as the θ_1 varies. The electronics are housed below

the arm in a box to raise the height of the arm above Temi's display.

The final design consideration comes in the form of setting the tension of the timing belts. Idler pulleys were added at the base to maintain the tension in the timing belts of the base motor. A groove is cut at the end of the base arm (at point B) to vary the tension of the belt driving the upper arm. Lastly, a elongated groove is cut on the motor mounts to tension the timing belts connecting the intermediate shafts with the motor.

All arm elements were manufactured on 5mm clear acryllic, while all elements on the base were manufactured on 8mm clear acryllic.

C. Computer Vision

We have three main goals with the Computer Vision aspect for the robotic arm: Patient Identification, Object Detection, Face Tracking, and Skin Anomaly Detection. For this paper, we limit our scope to traditional computer vision and do not evaluate machine learning and deep learning techniques.

Patient Identification

Each patient is identified through a QR code placed adjacent to the patient on the bed. This QR code contains information about the patient's name, identification number, primary medical practitioner, age, gender, etc. For the scope of this study this information is stored as plain text within the QR-code, however for future versions this information would be encrypted. We use the jsQR function to detect and decode the QR code, which is then interpreted by our web interface to display the patient information to the medical practitioner evaluating the patient.

Object Detection

Most traditional object detection mechanisms such as frame differencing and background subtraction have the implicit assumption of a stationary camera. That is, the model implicitly assumes the background does not move significantly in consecutive frames, which is not possible when the camera is moving.

Hence, we evaluate using Scale Invariant Feature Transform (SIFT) based features for object detection [6]. As these features are resilient to movement of the background within the frame, we can use the technique detailed in [7] for object detection.

Face Tracking

For the task of face-tracking we choose to use the Haar-Cascade algorithm [8]. This algorithm employs a machine learning approach based on a cascade of simple classifiers to achieve high computational efficiency and real-time performance. The detection process relies on three primary components: the integral image representation, the use of AdaBoost for feature selection, and a cascade structure for attentional filtering.

The algorithm does not operate directly on pixel intensities but rather on Haar-like features. These features compute the difference in summed pixel intensities between adjacent rectangular regions, effectively capturing local contrast variations characteristic of human facial structures (e.g., the intensity difference between the eye region and the cheek). To facilitate rapid feature evaluation, the input image is converted into an integral image representation. This intermediate representation allows for the calculation of the sum of pixels within any rectangular area in constant time regardless of the feature's scale or position.

Given the high dimensionality of the feature space, the AdaBoost learning algorithm is employed to select a small subset of critical visual features from a vast pool of potential candidates, yielding a strong classifier from a weighted combination of weak classifiers. To optimize processing speed, these classifiers are organized into a degenerate decision tree, or "cascade." The cascade functions as a sequential filter where early stages utilize a minimal number of features to reject the majority of non-face sub-windows with low computational cost (the "attentional cascade"). Only sub-windows that pass the initial stages with high probability are subjected to subsequent, more complex stages of classification, thereby focusing computational resources on promising regions of interest.

Skin Anomaly Detection

The skin anomaly detection pipeline employs a color-based segmentation approach in the HSV (Hue, Saturation, Value) color space, followed by geometric filtering to ensure robustness against false positives [9]. Initially, the input BGR image is converted to the HSV color space to decouple chromatic information from luminescence. This transformation is critical for isolating the target anomalies, which are defined by high brightness and low color saturation. A binary mask is generated by thresholding the HSV channels; pixels are retained only if they satisfy strict bounds for whiteness, specifically a Value range of $V \in [200, 255]$ and a Saturation range of $S \in [0, 50]$.

To mitigate noise and enhance the structural integrity of the detected regions, the binary mask undergoes morphological opening using a 3×3 kernel. This operation suppresses isolated noise pixels (salt-and-pepper noise) and separates weakly connected components. Following this refinement, the system extracts the external contours of the remaining connected components for geometric analysis. To differentiate valid anomalies from spurious detections, a series of geometric constraints are applied to the extracted contours. First, a spatial filter removes regions with an area A outside the range $100 < A < 1000$ pixels. Second, a shape factor analysis is performed using a circularity metric $\mathcal{C}$, defined as [10]:

$$C = \frac{4\pi A}{P^2} \tag{1}$$

where P represents the perimeter of the contour. Candidates with $C < 0.5$ are rejected to filter out elongated shapes that are inconsistent with the target morphology. Finally, the valid candidates are fitted with a bounding box based on contours.

We evaluate all our computer vision techniques based on their latency when deployed on the resource constrained

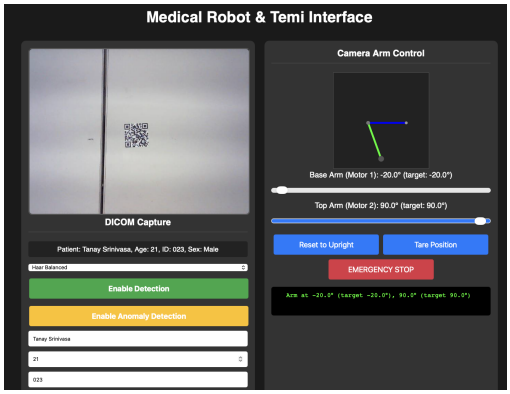

Fig. 3. User Interface: Computer Vision and Arm Control

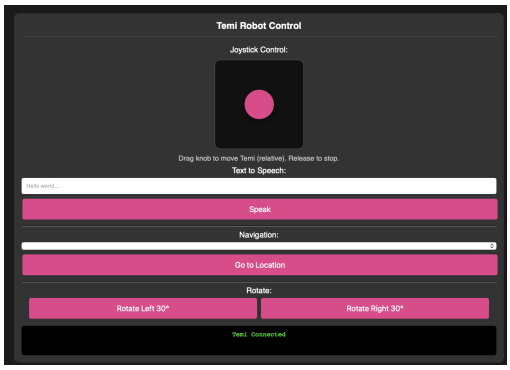

Fig. 4. User Interface: Temi Robot Control

Raspberry Pi 4. The comparison of their accuracy and efficacy is not within the scope of this paper.

D. Interface

The web interface, shown in Figs. 3 and 4, serves as the central control and visualization platform for the medical robot system, implemented using a Flask-based web server to enable real-time interaction via a browser. The system integrates computer vision for detection tasks, motor control for physical manipulation, and robot communication, all running on a Raspberry Pi 4B for portability and low-power operation. Future iterations of the medical robot system could address security requirements by implementing robust authentication mechanisms (e.g., OAuth or multi-factor authentication) and end-to-end encryption for data transmission and storage, ensuring compliance with standards like HIPAA for patient privacy.

III. Results

Robot Arm

Fig. 5 show the fully assembled robotic arm. As we observed a slight overshoot in the arm position during operation, we measured the settling time of oscillations for 200 steps for varied stepping times. The following results were observed:

From Table II we see that the settling time for the stepping time of 0.005s is the least. It was visually observed that the overshoot was the least for this stepping time. We noticed

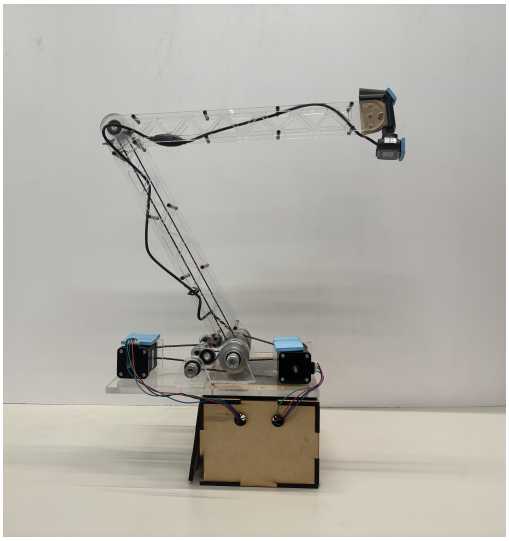

Fig. 5. Front View of Fully Assembled Robotic Arm

TABLE II
Settling Time for Oscillations for Varied Stepping Times

Stepping Time	Base Arm	Top Arm
0.005s	2.08s	2.25s
0.001s	3.08s	2.74s
0.0005s	3.01s	3.41s

smoother motion with lower overshoot for higher stepping times however the overall motion of the arm was too slow for final deployment. Hence the stepping time of 0.005s was chosen.

Computer Vision

The following latency results are from averaging inference times over 100 inferences.

TABLE III
Comparison of Inference Times of Various Computer Vision Techniques on the Raspberry Pi 4

Detection Method	Latency
Patient Identification	67.2 ms
SIFT	64.41 ms
Haar Cascade (Low Latency)	124.46 ms
Haar Cascade (Balanced)	203.63 ms
Haar Cascade (Accurate)	253.53 ms

As seen in Table III, SIFT based object detection is running at 15.5 frames per second, which is sufficiently real-time for such an application. Haar Cascade, based on accuracy, runs at 8, 4.9, or 3.9 frames per second. The low latency configuration of the Haar Cascade can run in real-time while dropping 2 out of every 3 frames at a video frame rate of 24 fps. More accurate configurations of the Haar cascade will encounter some buffered frames during deployment.

Skin Anomaly Detection

Fig. 6 demonstrates the working of the Skin Anomaly Detection technique explained in II-C. This demonstration was

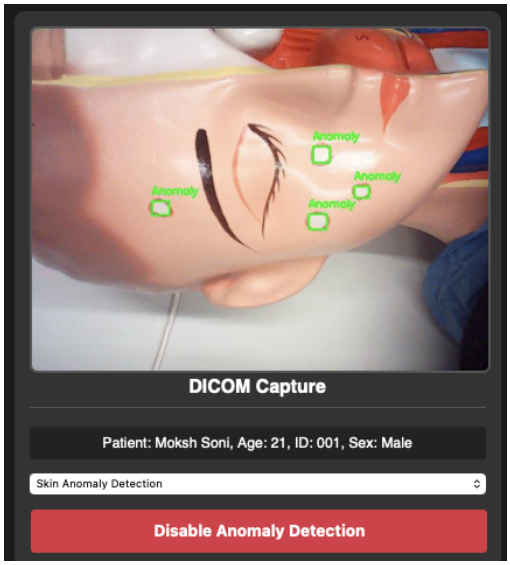

Fig. 6. Demonstration of Skin Anomaly Detection

conducted on a medical manequin to show feasibility of the technique.

IV. Discussions

Robot Arm

The Temi Robot has a maximum 3kg payload on its tray [11], hence the complete robot arm assembly must be within this weight. The current design uses a lightweight camera, which has a limited resolution. Discussions with surgeons highlighted low camera resolution as a primary concern. Higher resolution cameras are naturally heavier in nature, hence two changes would need to be implemented: further gear reduction in the motors to handle the weight; reduction in weight in the current assembly to accommodate the increased weight.

To tackle the increased torque load on the motor we propose that a gear reduction of 360:1 be implemented using a planetary gearbox on the motor shaft. This gear reduction would also reduce the jittering and overshoot currently experienced with the 6:1 gear reduction. To reduce the weight of the arm, we propose the removal of unused areas on the base of the arm. These regions currently are manufactured out of 8mm acrylic, a fairly heavy material. Removing these, would in part, help reduce the weight of the overall assembly.

Additionally, we propose moving the motors into the base box to abstract as many working elements on the arm from patients. As this arm would be operating in hospital environments, most working elements of the robot assembly, excluding the robot arm and camera, would need to be encapsulated into a single housing which cannot be accessed or tampered with by patients or unauthorized personnel.

Integration testing with the Robotic Arm mounted on Temi has not yet been carried out, and was not within the scope of this paper. Future work would include conducting this integration test. Currently, it has been observed that Temi's connectivity through its Java SDK is sensitive to the stability of the network connection. A momentary drop in the network connection requires a complete reboot of the entire system. We propose using a lower level communication with the Temi robot, directly with its internal ROS environment, to tackle this issue.

Access to Temi's internal ROS environment would also allow implementing a custom social navigation policy on Temi. Our tests show that Temi's navigation policy moves close to nearby moving obstacles at a quick speed which is not ideal for hospital environments. The internal ROS environment would give us access to internal maps and onboard sensor readings, allowing the implementation of a policy as defined in [12].

The final point of discussion with the surgeons led to the idea of adding UV sterilization equipment to the base of the robot, to help improve the decontaminate the hospital environment as the robot moves across the infectious ward.

Computer Vision

During testing with the Haar Cascade based face tracking, it was noticed that the technique was not resilient to tilting and turning of the head. That is, the face tracking would drop when the head was turned or tilted. Hence we propose the use of SIFT in conjunction to the Haar Cascade for face tracking. The Haar Cascade would serve as an initialization for the SIFT technique, after which SIFT based object tracking would be used.

The skin anomaly detector is highly sensitive to changes in illumination as it works on colour segmentation. Hence we propose the addition of a high intensity light source at the base of the camera, to reduce the effect of background lighting. However, the current technique is also sensitive to changes in skin colour. The use of machine learning based and deep learning based techniques could help with this issue.

Lastly, current experiments with the computer vision techniques do not measure efficacy, generalization and accuracy. Future work would contain collecting and labelling a dataset to quantify these metrics.

User Interface

The secure storage of data collected is subject to regulatory standards. The collected data would need to be transferred to a encrypted hospital server periodically, as well as periodic deletion of all locally saved scans and patient data to be in accordance to regulatory standards. Future work would also include understanding and ensuring compliance to such standards for the robot, computer vision technique, and user interface.

V. Conclusion

In this paper we designed and integrated a telepresence robotic arm for remote dermatological monitoring of infectious wards. This was motivated by the need to reduce the risk of nosocomial infections for susceptible patients and medical practitioners. For this system, we choose the Temi 3 robot with

a 2 degree-of-freedom robotic arm mounted on the Temi tray. We present design considerations, requirements, and optimizations of the Robot Arm. Additionally, we compare computer vision techniques for patient identification, object tracking, face tracking, and skin anomaly detection. We benchmark these techniques by measuring and comparing their inference latency to test real-time deployment feasibility. Lastly we present considerations and discussions for future versions of this system.

REFERENCES

[1] World Health Organization, *Global report on infection prevention and control 2024*. Geneva: World Health Organization, 2024. [Online]. Available: https://www.who.int/publications/i/item/9789240103986

[2] European Centre for Disease Prevention and Control, *Point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals 2022-2023*. Stockholm: ECDC, 2024.

[3] H. Melariri, R. Freercks, E. van der Merwe, W. T. Ham-Baloyi, O. Oyedele, R. A. Murphy, C. Claasen, P. E. Etusim, M. O. Achebe, S. Offiah, and P. E. Melariri, "The burden of hospital-acquired infections (hai) in sub-saharan africa: a systematic review and meta-analysis," *eClinicalMedicine*, vol. 71, 2025/12/06 2024. [Online]. Available: https://doi.org/10.1016/j.eclinm.2024.102571

[4] O. Rasslan, A. El-Kholy, E. El-Sheshtawy *et al.*, "Incidence of hospital-acquired infections among healthcare workers in egypt before and during the covid-19 pandemic," *Eastern Mediterranean Health Journal*, vol. 30, no. 9, pp. 612–621, 2024.

[5] M. Haque, M. Sartelli, J. McKimm, and M. A. Bakar, "Health care-associated infections – an overview," *Infection and Drug Resistance*, vol. 11, pp. 2321–2333, 2018.

[6] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004. [Online]. Available: https://doi.org/10.1023/B:VISI.0000029664.99615.94

[7] S. Gu, Y. Zheng, and C. Tomasi, "Efficient visual object tracking with online nearest neighbor classifier," in *Computer Vision – ACCV 2010*, R. Kimmel, R. Klette, and A. Sugimoto, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 271–282.

[8] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 1, 2001, pp. I–I.

[9] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat, and J. Jatakia, "Human skin detection using rgb, hsv and ycbcr color models," in *International Conference on Communication and Signal Processing 2016 (ICCASP 2016)*, 08 2017.

[10] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and M. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320314000235

[11] Temi USA Inc., *Temi Robot Specifications*, 2025, accessed: 2025-12-07. [Online]. Available: https://www.robotemi.com/product/temi/

[12] R. Malladi, A. Harsh, A. Sangwan, S. Chauhan, and S. Manjanna, "Sango: Socially aware navigation through grouped obstacles," in *2024 Tenth Indian Control Conference (ICC)*, 2024, pp. 526–531.

APPENDIX

A. Torque Calculation

This section presents the estimated torque requirements of the bottom and top arms in their extreme positions. The required torques are compared with the motor torques after gear reduction. All lengths and masses are referenced from Fig. 7.

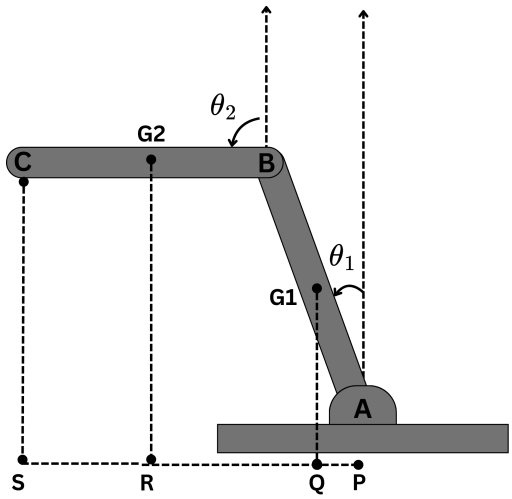

Fig. 7. Extreme Configuration of the 2-DoF Robotic Arm

Variable Definitions: θ_1, θ_2 – Joint angles; PQ, PR, PS – Distances from bottom axis to CG of bottom arm, CG of top arm, and camera; QR, QS – Distances from top axis to CG of top arm and camera; W_b, W_t, W_c – Weights of bottom arm, top arm, and camera; τ_m – Rated motor torque; G_b, G_t – Gear ratios; $\tau_{\text{req},b}, \tau_{\text{req},t}$ – Required torques; $\tau_{\text{avail},b}, \tau_{\text{avail},t}$ – Available torques after gear reduction.

$$\theta_1 = 20^\circ, \qquad \theta_2 = 90^\circ$$

Geometric Distances:

$$PQ = AG_1 \sin\theta_1 = 5.98 \text{ cm}$$
$$PR = 2PQ + BG_2 \sin\theta_2 = 29.47 \text{ cm}$$
$$PS = AB \sin\theta_2 + BC \sin\theta_2 = 46.97 \text{ cm}$$
$$QR = BG_2 \sin\theta_2 = 17.5 \text{ cm}$$
$$QS = BC \sin\theta_2 = 35 \text{ cm}$$

Weights: $W_b = 0.14$ kg, $W_t = 0.12$ kg, $W_c = 0.10$ kg
Motor Parameters: $\tau_m = 6.5$ kgcm, $G_b = 6$, $G_t = 3$

Torque Requirements:

$$\tau_{\text{req},b} = PQ\,W_b + PR\,W_t + PS\,W_c = 9.07 \text{ kgcm},$$
$$\tau_{\text{avail},b} = \tau_m G_b = 39 \text{ kgcm},$$
$$\tau_{\text{req},t} = QR\,W_t + QS\,W_c = 5.6 \text{ kgcm},$$
$$\tau_{\text{avail},t} = \tau_m G_t = 19.5 \text{ kgcm}.$$

Both motors provide sufficient torque for the given extreme configuration.