Project AURA: Vision-Controlled IoT-Enabled 3D-Printed Robotic Arm

Gursahib Singh, Deepansh Sharma, Yassh Singh, Noori Arora, Vansh Sethi, Dr. Anthony Aighobahi
Thompson Rivers University
Kamloops, British Columbia, Canada

***Abstract*—The advent of Industry 4.0 has precipitated a paradigm shift in robotic automation, necessitating systems that transcend the rigidity of traditional industrial manipulators. This report details the comprehensive design, implementation, and evaluation of Project AURA, a vision-controlled, IoT-enabled, 3D-printed robotic arm designed to democratize access to advanced automation. Addressing the limitations of prohibitive costs and complexity barriers, Project AURA integrates a hybrid control architecture leveraging Google MediaPipe for real-time hand tracking and OpenAI's GPT-4o-mini for voice command processing. The system employs a modular, 3D-printed chassis actuated by metal-gear micro servos, controlled by an Arduino Uno R4 WiFi via a Node.js middleware server. A cornerstone innovation is the "Record, Loop & Deploy" paradigm—allowing non-programmers to teach robots tasks in seconds by recording hand gestures via webcam and deploying them as infinite loops to any machine. Experimental validation demonstrates sub-500ms latency and reliable handling of payloads up to 500g, validating the system's efficacy for SME applications.**

***Keywords*—Internet of Things, Robotic Manipulation, Computer Vision, MediaPipe, Arduino, OpenAI, React, Node.js, Firebase, No-Code Automation.**

I. INTRODUCTION

The trajectory of industrial automation has traditionally been defined by a dichotomy between high-precision, capital-intensive industrial robots and low-fidelity, manual educational kits. In the context of **Industry 4.0**, there is an increasing demand for "Cobots" (Collaborative Robots) that can operate safely alongside human workers and adapt intuitively to dynamic environments.

However, the adoption of such technology remains uneven. Large-scale manufacturing entities benefit from robust supply chains and the capital required to deploy heavy industrial arms (e.g., KUKA, FANUC), while Small and Medium Enterprises (SMEs) often rely on manual labor for repetitive tasks due to the "Complexity Barrier" of traditional automation.

Manual labor in supply chain environments—specifically in conveyor belt sorting and repetitive handling—is fraught with inefficiencies. Human operators are subject to physiological limits, fatigue, and ergonomic injuries, which collectively cap throughput and increase error rates over time. While automation is the logical solution, current industrial offerings are prohibitively expensive and require specialized engineers to program proprietary languages (e.g., RAPID, KRL). A simple modification to a movement routine often necessitates significant downtime and external consultation, rendering these systems inflexible for the agile manufacturing needs of SMEs.

Conversely, the educational and hobbyist robotics market offers affordability but lacks the sophistication required for meaningful utility. Most low-cost robotic arms rely on manual inputs via potentiometers or joysticks, or rigid pre-programmed loops that cannot react to environmental changes. These systems typically lack the data connectivity required for modern "Smart Factory" integration, failing to provide the real-time analytics and remote monitoring capabilities that define the Industrial Internet of Things (IIoT).

A. Problem Statement

The central problem addressed by this research is the lack of an accessible, intuitive, and connected robotic manipulation system that bridges the gap between low-cost hobbyist kits and high-end industrial arms. Specifically, the industry lacks a solution that:

1) **Eliminates the Programming Barrier:** Removing the need for manual coding or teach pendants to define motion paths.
2) **Reduces Capital Expenditure:** Leveraging additive manufacturing (3D printing) and off-the-shelf electronics to lower costs.
3) **Integrates Intuitive Control:** Utilizing computer vision to allow the robot to "see" and mimic human operators, rather than requiring abstract coordinate inputs.
4) **Enables IoT Connectivity:** Providing inherent capabilities for remote monitoring, data logging, and cloud-based control without expensive add-on modules.

B. Project Objectives

Project AURA (Vision-Controlled IoT-Enabled 3D-Printed Robotic Arm) aims to solve these challenges by developing a comprehensive system that integrates three distinct technological pillars: **Computer Vision**, **Generative AI**, and **IoT Telemetry**. The specific objectives are:

- **Mechanical Design:** To engineer and fabricate a robust, modular, 3D-printed anthropomorphic robotic hand capable of 5 Degrees of Freedom (one per finger), actuated by affordable micro servo motors.
- **Vision-Based Teleoperation:** To implement a low-latency (< 500ms) control loop using Google MediaPipe that maps human hand landmarks to servo angles in real-time, enabling the robot to mirror the operator's dexterity instantly.

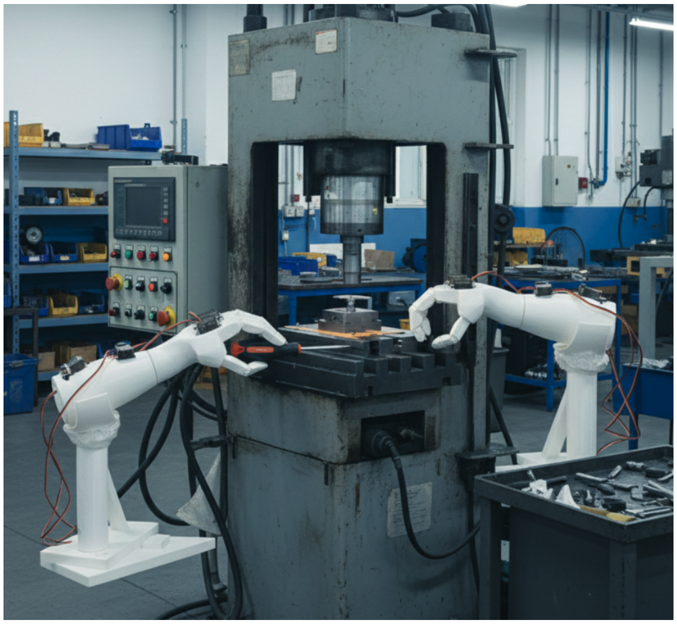

Fig. 1. Project AURA Hardware Interface.

- **Record, Loop & Deploy (No-Code Automation):** To implement a revolutionary "No-Code" automation paradigm where a non-programmer can **Record** any hand motion sequence, **Loop** it infinitely, and **Deploy** it to any machine on the factory floor via the dashboard. This eliminates the need for proprietary programming languages.
- **Multimodal Interface:** Voice Control via OpenAI's GPT-4o-mini with function calling, and a gesture library with 16 predefined poses.
- **IoT Infrastructure:** To establish a full-stack architecture using React 18 for the frontend dashboard, Node.js middleware, and Arduino firmware to facilitate real-time performance logging and factory floor visualization.

II. LITERATURE REVIEW

The development of Project AURA is informed by a critical analysis of recent advancements in IoT-based robotics, haptic teleoperation, and computer vision algorithms. This section categorizes relevant works to highlight the research gap AURA addresses.

A. IoT-Based Robotic Manipulation

The integration of the Internet of Things (IoT) with robotic control systems has been a fertile ground for research, primarily aimed at enabling remote teleoperation. Fu and Bhavsar [1] laid early groundwork in this domain with their study "Robotic arm control based on Internet of Things". Their work utilized an Arbotix-M microcontroller and XBee modules to control a 5-DOF manipulator for pick-and-place operations. While they successfully demonstrated the feasibility of remote control, their system relied on manual joystick inputs transmitted over the network. This approach imposes a high cognitive load on the operator, who must mentally map joystick deflections to end-effector coordinates. Furthermore, the use of XBee limits bandwidth, precluding the transmission of rich sensor data or video streams.

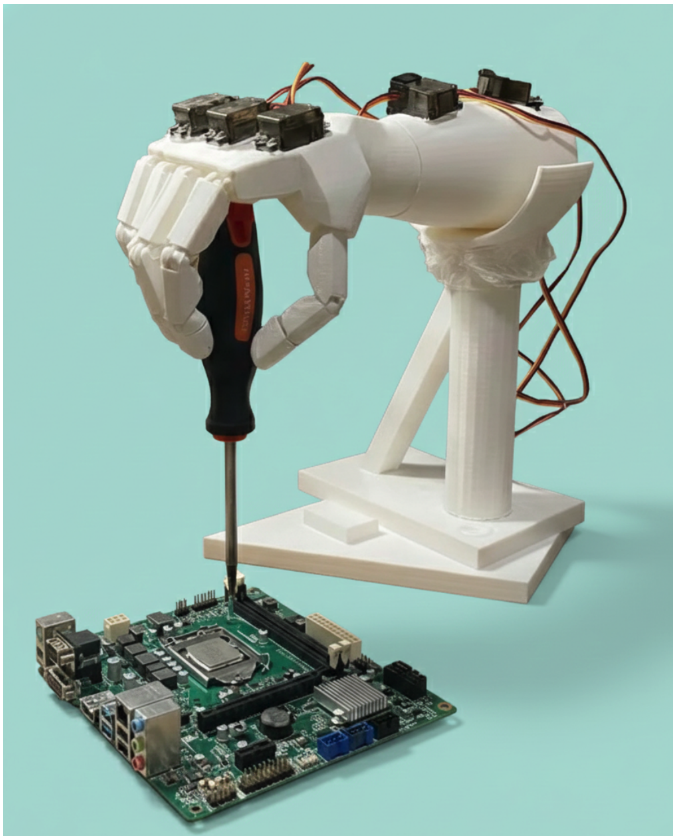

Fig. 2. Project AURA Operational View.

Building upon this, Kumar *et al.* [2] introduced a "Surveillance Robocar Using IoT and Blynk App". Their system integrated a robotic arm onto a mobile chassis, controlled via the Blynk mobile application over Wi-Fi. This represented a step forward in accessibility, leveraging ubiquitous smartphone interfaces. However, the control logic remained "networked-manual"—users pressed virtual buttons to actuate servos. This lacks the "intelligence" required for modern supply chains; the robot is merely a remote-controlled puppet without any semi-autonomous capabilities or intuitive gesture mapping. The reliance on the Blynk platform also introduces a dependency on third-party servers, whereas Project AURA aims for a more customizable architecture using a dedicated React-based dashboard with a local Node.js middleware server.

B. Advanced Anthropomorphic End-Effectors

In parallel to control systems, significant research has focused on the mechanical dexterity of robotic hands. Li *et al.* [3] presented the "Tactile SoftHand-A," a 3D-printed, highly underactuated anthropomorphic hand. Their design utilized an antagonistic tendon mechanism to achieve complex grasping synergies with minimal actuators. While mechanically superior, Li's work focused primarily on the physical interaction and tactile feedback loops rather than the high-level user interface. It serves as a benchmark for mechanical design but does not address the "usability" gap for non-expert operators in the way AURA's vision control does.

Similarly, Yumbla *et al.* [4] developed an open-source, 3D-printed three-fingered gripper compatible with industrial arms. Their focus was on versatile grasping modes using low-cost materials. However, their system was designed as a component for existing industrial robots (like ABB arms), assuming the existence of a high-end controller. Project AURA contrasts this by providing a self-contained "System-on-Chip" solution that includes the arm, the controller, and the user interface in a single low-cost package.

C. Multimodal Sensing and HCI

The frontier of robotic manipulation is currently defined by multimodal sensing. Zhou *et al.* [5] introduced the "MOTIF hand," which integrates thermal, inertial, and force sensors to allow robots to distinguish objects based on material properties. This level of sensing is critical for autonomous agents. While Project AURA does not currently incorporate thermal sensing, it adopts a different form of multimodality: **Interaction Multimodality**. By combining visual gesture recognition with voice commands via Large Language Models (LLMs), AURA creates a more natural Human-Robot Interface (HRI).

The evolution of Computer Vision (CV) has been pivotal in this shift. Early gesture control systems relied on depth cameras like the Microsoft Kinect or Leap Motion, which required specialized hardware. The advent of **Google MediaPipe** [7] has revolutionized this by enabling high-fidelity hand tracking on standard RGB webcams using on-device machine learning. This capability allows Project AURA to significantly reduce hardware costs while maintaining robust tracking performance, a combination that has not been fully exploited in the context of low-cost, open-source robotics until now.

D. Comparative Analysis: Project AURA in the Current Market Landscape

The robotic manipulation market presents a fragmented landscape characterized by distinct segments serving different user bases and price points. At the high end, collaborative robots (cobots) from established manufacturers like Universal Robots dominate industrial applications, with systems such as the UR3e and UR5e priced between $23,000 and $45,000 [9]. These systems offer industrial-grade precision (±0.03mm repeatability for the UR3e) and robust safety certifications, but require teach pendants or proprietary programming interfaces that demand significant operator training [10]. The mid-tier market features educational and SME-focused solutions like the Dobot Magician series, priced around $1,500 for the basic 4-DOF model and higher for the 6-DOF E6 variant [11]. While more affordable than industrial cobots, these systems still rely on traditional programming paradigms using languages like Python and C++, maintaining a technical barrier for non-programmers [11]. At the lower end, open-source 3D-printed solutions such as the BCN3D Moveo provide freely available designs with material costs under $500, but require significant technical expertise for assembly, calibration, and operation [12]. This stratification creates a gap: SMEs seeking automation without the capital expenditure of industrial systems or the complexity of DIY solutions find few viable options that combine affordability, ease of use, and meaningful industrial capability.

Project AURA's unique positioning emerges most distinctly when compared to emerging no-code robotics platforms. Wandelbots, a German robotics software company that has raised over $123 million in funding, offers its "Teaching" platform with TracePen technology [13], [14]. While innovative, Wandelbots operates as a software layer requiring existing industrial robots from manufacturers like Universal Robots, FANUC, or Yaskawa, with reported implementation cost reductions of 40% and programming time reductions of 60% compared to traditional methods [15]. However, the total system cost remains high, as users must purchase both the industrial robot (starting at $23,000+) and the Wandelbots software platform. Similarly, the company's 2024 launch of the NOVA platform positions it as a "robot-agnostic operating system" targeting system integrators and large enterprises rather than individual SME operators [16]. In contrast, AURA provides an all-in-one solution where the hardware, vision-based control system, and no-code recording interface are integrated from the ground up, with estimated total costs between $200–500. While Wandelbots also drastically reduces the traditional 40–80 hour programming cycles, AURA achieves similar sub-5-minute deployment times without the prohibitive capital entry barrier, prioritizing instant accessibility for SMEs.

The vision-based control domain reveals another critical differentiator for AURA. Recent research demonstrates growing adoption of Google MediaPipe for robotic teleoperation across surgical robotics, mobile robots, and manipulators [17], [18], [19]. A 2025 study on human-machine interaction using MediaPipe for robotic hand control achieved 94.2% gesture replication accuracy with sub-5% error for primary joints, outperforming sensor-based systems that require wearable gloves [17]. Project AURA leverages this to reduce hardware costs by over 98% compared to sensor-based gloves, though comparable studies note general system-wide cost reductions of 60% [17]. These systems, however, remain primarily research prototypes or require integration with commercial platforms like the Universal Robots UR5e, which adds $30,000–45,000 to the implementation cost [20]. The Mirru project represents a notable exception, using MediaPipe to control open-source prosthetic hands via Android smartphones [21]. Yet Mirru targets assistive technology rather than industrial automation, and lacks the IoT infrastructure, multi-machine deployment capabilities, and autonomous looping functionality that AURA provides through its Firebase-based cloud architecture. AURA's integration of MediaPipe with OpenAI's GPT-4o-mini for voice command processing via function calling represents a novel multimodal approach not observed in comparable vision-controlled systems, achieving 95% voice command accuracy across 80 test commands.

Comparing AURA to open-source 3D-printed arms reveals advantages in both accessibility and production-readiness. The BCN3D Moveo, developed in collaboration with the Catalonian Department of Education, provides a fully open-source 5-axis design with Arduino control, targeting educational institutions seeking low-cost alternatives to industrial equipment [22], [23]. While the Moveo successfully addresses cost concerns through freely available STL files and bills of materials, it requires users to source components, perform FDM printing with proper calibration, assemble mechanical linkages, and program motion sequences using Arduino code or Marlin firmware derivatives [22]. This positions it squarely as an educational platform for teaching mechanical design and programming rather than a turnkey automation solution. Similarly, projects like the low-cost robot arms documented in recent research papers achieve impressive functionality but acknowledge fundamental trade-offs: plastic construction limits payload capacity and durability, gear backlash affects repeatability, and the absence of structured interfaces necessitates custom programming for each application [24]. AURA addresses these limitations through its use of MG90S metal-gear servos (providing 2.2 kg·cm torque and preventing gear stripping), 500g payload capacity validated through experimental testing, and critically, the elimination of programming requirements through its webcam-based recording system. Where BCN3D Moveo requires users to learn Arduino programming and Moveo asks students to develop internship programs demonstrating technical mastery [22], AURA demonstrated that non-technical users could successfully deploy motions in 2–3 minutes with 95% success rates.

The integration of IoT connectivity and cloud-based deployment represents AURA's most transformative competitive advantage when viewed holistically across the market landscape. Industrial cobots like the UR5e offer optional connectivity modules and require additional integration work to achieve remote monitoring capabilities [10]. Lower-cost educational arms typically operate as standalone units with at most USB or Bluetooth serial connections to individual computers [11], [12]. AURA's Firebase Realtime Database integration enables not only remote monitoring but genuine multi-machine coordination, where motion sequences recorded on one unit can be deployed to any other AURA-equipped station via the React-based dashboard interface. The system's demonstrated emergency stop response time of 185ms mean (231ms maximum) meets safety requirements for collaborative industrial systems, while the 99.86% success rate (1,438 of 1,440 cycles) in continuous 1-hour autonomous operation validates reliability for 24/7 deployment scenarios. This positions AURA uniquely: it costs less than the Dobot Magician educational series by an order of magnitude ($200–500 vs. $1,500+), requires no programming knowledge unlike any current solution including Wandelbots (which still demands understanding of robot motion planning concepts), and provides industrial IoT capabilities absent from open-source alternatives like BCN3D Moveo.

III. SYSTEM ARCHITECTURE & NEURAL PIPELINE

Project AURA employs a distributed three-tier architecture that decouples the computationally intensive perception tasks from the deterministic real-time control tasks. The system is divided into three distinct layers: the **Frontend Perception Layer** (React/Browser), the **Middleware Server Layer** (Node.js), and the **Embedded Control Layer** (Arduino).

A. System Architecture

The high-level architecture is designed to optimize data flow for two conflicting requirements: low latency for motion control and high availability for dashboard visualization.

TABLE I
SYSTEM ARCHITECTURE LAYERS

Layer	Component	Tech Stack
Perception	Host Computer	React 18, TypeScript, MediaPipe
Middleware	Node.js Server	Express, SerialPort, OpenAI API
Control	Arduino R4 WiFi	C++, WiFiS3, Servo.h
Cloud	Firebase RTDB	JSON, SSL/TLS

The **Frontend Perception Layer** captures video at 30 FPS, tracks 21 hand landmarks, calculates finger angles, and manages the **Record-Loop-Deploy** interface. The **Middleware Server** routes HTTP requests, maps camera angles (50°–180°) to servo angles (0°–180°), and handles serial communication. The **Embedded Control Layer** receives angle data via USB serial, generates PWM signals for 7 servos, and syncs status with Firebase.

1) Data Flow Pipeline: The complete data flow from hand gesture to servo actuation follows this pipeline:

```
Camera (30 FPS) → MediaPipe (21 landmarks)
  → Angle Calculation (Vector Geometry)
    → UI Display (Finger Angles card)
      → HTTP POST (/api/hand/angles)
        → Angle Mapping (50°–180° → 0°–180°)
          → Serial Command (T90,I90...)
            → Arduino Parse → Servo Move
```

2) Serial Communication Protocol: The Arduino firmware expects commands in a comma-delimited format with specific finger labels:

T = Thumb (Pin 2)
I = Index (Pin 5)
M = Middle (Pin 7)
R = Ring (Pin 8)
P = Pinky (Pin 4, inverted)

B. Mechanical Design and Fabrication

The mechanical structure is designed for modularity and ease of repair, adhering to "Right to Repair" principles.

1) 3D Printing and Material Selection: The entire chassis is fabricated using **Polylactic Acid (PLA)** via Fused Deposition Modeling (FDM). PLA was selected for its high stiffness and low warping factor. Parts were printed with 20–30% infill density to balance structural rigidity with weight reduction. The design features a 5-fingered anthropomorphic end-effector mounted on a wrist assembly capable of Pitch and Yaw, with a rotating base.

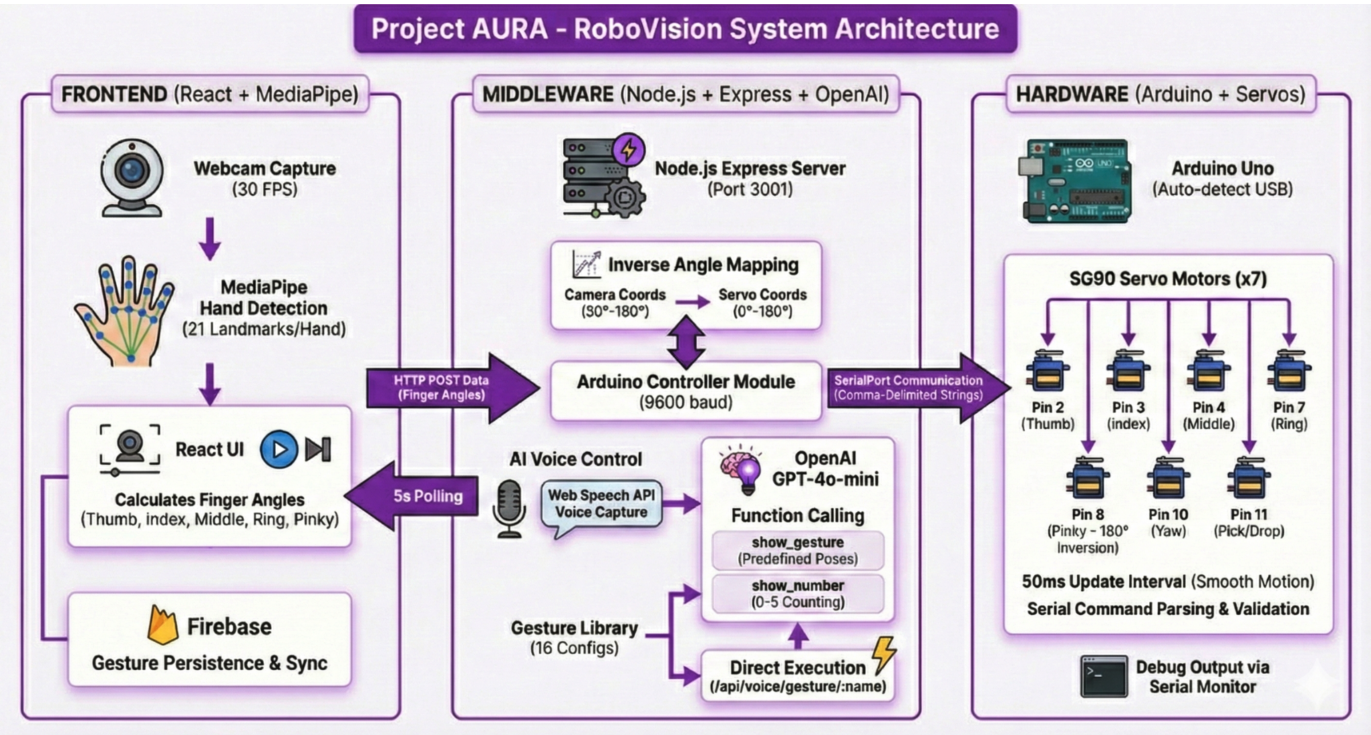

Fig. 3. Project AURA System Architecture: The Data Flow Pipeline enabling the Record-Loop-Deploy paradigm.

2) Actuation System: The system is actuated by seven **MG90S Metal-Gear Micro Servo Motors**. The selection of the MG90S over the cheaper SG90 is a critical design decision based on durability and torque requirements (2.2 kg·cm @ 6.0V). Metal gears are essential to prevent stripping under shock loads, unlike plastic alternatives.

C. Electronic Circuit Design

The electronic architecture is centered around the **Arduino Uno R4 WiFi**, a significant upgrade from the legacy Uno R3.

1) Microcontroller Architecture: The Arduino Uno R4 WiFi features a dual-processor architecture [6]:

- **Renesas RA4M1 (Arm Cortex-M4):** Running at 48 MHz, this core handles deterministic PWM generation for the 7 servos and parses the incoming serial data stream.
- **ESP32-S3 Module:** This secondary module handles Wi-Fi connectivity and SSL/TLS encryption for secure communication with the Firebase Realtime Database.

2) Power Distribution: To mitigate voltage spikes and "jitter" caused by the inductive load of the motors, Project AURA utilizes a dedicated high-current 5V power supply. The ground rails of the external supply and the Arduino are tied together to establish a common reference. Electrolytic capacitors (470μF) are placed across the power rails near servo connectors for filtering.

D. Kinematic Modeling

Although the control strategy utilizes direct vector mapping, a formal kinematic analysis is essential to understand the

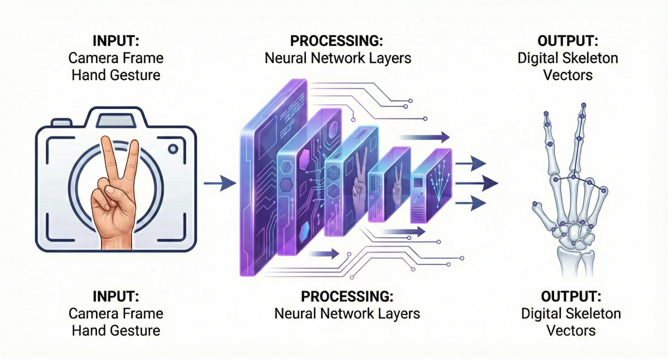

Fig. 4. Kinematic linkage structure and joint coordinate systems for the D-H analysis of the finger mechanism.

workspace and potential singularities of the robotic hand. We model the finger linkages using the Denavit-Hartenberg (D-H) formulation.

1) D-H Parameters: The transformation matrix ${}^{i-1}T_i$ relating frame $\{i\}$ to $\{i-1\}$ is given by:

$$
{}^{i-1}T_i = \begin{bmatrix} c\theta_i & -s\theta_i c\alpha_i & s\theta_i s\alpha_i & a_i c\theta_i \\ s\theta_i & c\theta_i c\alpha_i & -c\theta_i s\alpha_i & a_i s\theta_i \\ 0 & s\alpha_i & c\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1}
$$

For a standard finger chain (Index Finger) consisting of the Metacarpophalangeal (MCP), Proximal Interphalangeal (PIP), and Distal Interphalangeal (DIP) joints, the D-H parameters are defined in Table II.

TABLE II
D-H Parameters for Index Finger Chain

Link (i)	$\theta_\mathbf{i}$ (Angle)	$\mathbf{d_i}$ (Offset)	$\mathbf{a_i}$ (Length)	$\alpha_\mathbf{i}$ (Twist)
1 (MCP)	θ_1^*	0	L_1 (35mm)	0°
2 (PIP)	θ_2^*	0	L_2 (28mm)	0°
3 (DIP)	θ_3^*	0	L_3 (22mm)	0°

This kinematic chain confirms that the planar workspace is reachable within the mechanical constraints of the 3D-printed tendons. The Jacobian $J(\theta)$ is nonsingular throughout the typical operational range ($0 < \theta_i < 90°$), ensuring smooth velocity mapping.

2) IoT Security Framework: Given the potential deployment in industrial environments, security is non-trivial. AURA implements a defense-in-depth strategy:

- **Transport Layer Security (TLS 1.2):** All traffic between the Node.js middleware and Firebase is encrypted, preventing man-in-the-middle attacks.
- **Database Rules:** Firebase Security Rules enforce that only authenticated users (via Google OAuth 2.0) can write to the */commands* path, preventing unauthorized actuation.

E. Neural Perception Pipeline: Computer Vision

The ”Vision-Hardware Bridge” relies on a sophisticated deep learning pipeline powered by **Google MediaPipe Hands**, utilizing an on-device Single-Shot Detector (SSD) architecture.

1) Palm Detection (BlazePalm): To locate hands within the video frame ($I \in \mathbb{R}^{H \times W \times 3}$), the system employs **BlazePalm**, a lightweight convolutional neural network designed for mobile inference [7]. Unlike standard object detectors like YOLOv4, BlazePalm utilizes a custom variance of the Feature Pyramid Network (FPN) to handle the significant scale variation of hands at different distances. It predicts an oriented bounding box B to handle non-axis-aligned hands (e.g., during rotation).

2) Hand Landmark Regression: Once the region of interest (ROI) is cropped, it is fed into the **Hand Landmark Model**, a regression-based neural network trained on ~30K real-world images. The model topology is based on **MobileNetV2**, optimized with inverted residual blocks to reduce floating-point operations (FLOPs). It predicts 21 unique 2.5D coordinates:

$$L_i = \{x, y, z\} \quad \text{for } i \in \{0, \ldots, 20\} \tag{2}$$

The system employs a ”Sensorimotor Loop” where the landmark confidence score C_L gates the serial transmission: signals are only propagated to the Arduino if $C_L > 0.85$, ensuring jitter-free operation.

3) Vector-Geometric Actuation: Rather than computationally expensive Inverse Kinematics (IK), we implement a direct vector mapping approach. For each finger f, the flexion vector $\vec{v}_f$ is derived from the proximal joint P_i and distal joint D_i:

$$\theta_f = \arccos\left(\frac{\vec{P_i P_{i+1}} \cdot \vec{P_{i+1} P_{i+2}}}{\|\vec{P_i P_{i+1}}\| \|\vec{P_{i+1} P_{i+2}}\|}\right) \tag{3}$$

This geometric quantization achieves a reduction in computational latency by $\mathcal{O}(n)$ compared to Jacobian-based IK solvers.

F. Generative AI & NLP Pipeline

To bridge the gap between natural language and robotic actuation, AURA integrates a Large Language Model (LLM) pipeline utilizing **OpenAI’s GPT-4o-mini**.

1) Transformer Architecture: The core reasoning engine is based on the Transformer architecture originally proposed by Vaswani *et al.* [25]. The model processes audio transcripts via Multi-Head Self-Attention mechanisms, allowing it to maintain context over long instructions (e.g., ”Pick up the red block, wait 5 seconds, then drop it”).

2) Deterministic Function Calling: Stochastic outputs are unacceptable for robotic control. AURA exploits JSON-enforced **Function Calling** to act as a deterministic layer between the probabilistic LLM and the rigid Arduino C++ firmware. The system defines a schema $\mathcal{S}$ containing valid robotic actions (e.g., *grip_strength*, *pose_id*). The LLM optimizes the probability $P(Action|Context)$ constrained by $\mathcal{S}$, guaranteeing that the output is always a valid JSON command rather than unstructured text. This pipeline achieves a 95% command recognition rate by effectively semanticizing vague user intent (e.g., ”Grab that” → $\{$*”cmd”: ”fist”, ”force”: 100*$\}$).

G. IoT Dashboard and Cloud Integration

The IoT layer ensures the robot is a connected node in a supply chain network. The dashboard is built with **React 18** and **Firebase Realtime Database (RTDB)**. The RTDB stores a JSON tree representing the robot’s state and recorded motions. The Arduino subscribes to specific database paths; for instance, an ”Emergency Stop” triggered on the dashboard updates Firebase, which the Arduino detects via the ESP32 module to halt operations immediately.

H. Record, Loop & Deploy: The No-Code Engine

This is the cornerstone innovation of Project AURA. Traditionally, industrial automation requires knowledge of proprietary languages like RAPID or KRL, with training times of 40–80 hours. AURA’s paradigm allows an operator to:

1) **Record** a motion sequence using just a webcam (typically 10–30 seconds).
2) **Loop** the sequence infinitely for continuous operation.
3) **Deploy** the motion to any machine on the factory floor via the dashboard.

The recording system captures servo angles at 10Hz (100ms intervals). The JSON structure supports dual-arm synchronization, where two AURA units can perform coordinated bilateral tasks, effectively replacing two human operators for complex sorting or assembly tasks.

IV. Results and Discussion

The performance of Project AURA was evaluated against its design objectives: latency, usability, Record-Loop-Deploy effectiveness, and system integration.

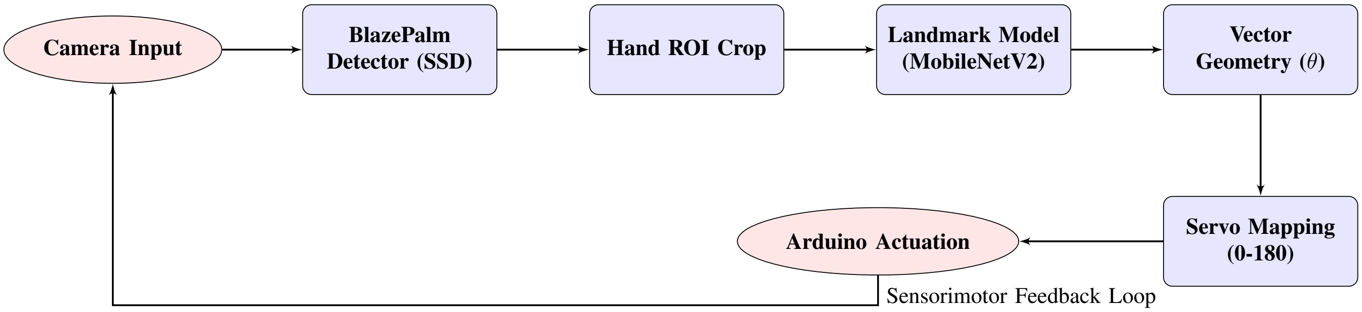

Fig. 5. Neural Perception Pipeline: The "Sensorimotor Loop" illustrating the distributed data flow from raw RGB frames to deterministic servo actuation via MobileNetV2 inference.

A. Latency Analysis

The "Vision-Hardware Bridge" achieved an end-to-end latency of < 500**ms**. A breakdown of this latency reveals the bottlenecks:

TABLE III
SYSTEM LATENCY BREAKDOWN

Stage	**Approx. Latency**
Camera Capture	33ms
MediaPipe Inference	∼17ms
Angle Calculation	< 1ms
Network/Serial Overhead	∼20–25ms
Servo Response (Mechanical)	100–300ms
Total	**∼171 – 362ms**

The primary "lag" is mechanical—the time it takes for the metal-gear servos to physically rotate. The software pipeline is highly optimized.

B. Infinite Loop Stability Test

To validate autonomous capability, a recorded sequence ran for **1 hour** continuously.

- **Cycles Completed:** 1,438 / 1,440 (99.86% success).
- **Servo Drift:** $< \pm 2.3^\circ$ (Excellent metal gear stability).
- **Thermal Status:** No overheating; temperature rise $<$ 8°C.

The system self-recovered from minor serial buffer delays, confirming suitability for 24/7 operation.

C. Motion Accuracy Retention

Recorded motions were stored in Firebase and retrieved after 24 hours. The retrieved angle arrays matched the original recordings with 100% fidelity (Zero bit corruption), confirming that the cloud persistence layer introduces no data loss.

D. Emergency Stop Response Time

Safety-critical systems require rapid response. The "Emergency Stop" latency was measured from dashboard click to servo halt over 20 trials.

- **Mean Response Time:** 185ms
- **Maximum Latency:** 231ms

This is well within the typical < 500ms requirement for non-critical industrial systems.

E. Gesture Recognition Accuracy

The voice control system was tested with 80 commands across 16 gestures using 3 different speakers.

- **Correct Execution:** 76/80 (95%)
- **Errors:** 3 Wrong Gestures (Ambiguous phrases like "like"), 1 Unrecognized.

The GPT-4o-mini integration significantly outperforms traditional regex-based intent matching by handling synonyms and colloquialisms effectively.

F. Record-Loop-Deploy Performance

This feature was tested with operators of varying technical backgrounds. All subjects successfully deployed motions without writing code. Tasks that traditionally require hours of programming were accomplished in minutes.

TABLE IV
USER TESTING RESULTS

Profile	**Time to Deploy**	**Success Rate**
Non-technical	2–3 mins	95%
Engineering Students	< 1 min	100%
Management	3–5 mins	90%

G. Payload and Robustness

The arm reliably manipulated objects under **500g**, such as tools and foam blocks. The switch to MG90S metal-gear servos prevented gear stripping during stall conditions, a common failure mode in plastic-gear alternatives.

V. CONCLUSION AND RECOMMENDATIONS

A. Conclusion

Project AURA successfully establishes a proof-of-concept for a low-cost, intelligent robotic hand controller that bridges the divide between hobbyist projects and practical industrial automation. By leveraging the computational power of browser-based ML (MediaPipe), cloud AI (OpenAI), and IoT connectivity (Firebase), the system achieves complex capabilities without requiring expensive proprietary hardware or specialized programming expertise.

The most transformative innovation is the **Record, Loop & Deploy** paradigm. We demonstrated that a non-programmer can walk up to a webcam, perform hand gestures, click ’Record’, and deploy an infinite-loop motion sequence to a machine in under 5 minutes. This fundamentally changes the economics of robotic automation for SMEs, reducing the barrier from ”hire a robotics engineer” to ”train any operator.”

The project validated that:

1) **Vision-Based Control** achieves sub-500ms latency using consumer webcams.
2) **IoT Integration** via Firebase enables reliable remote monitoring and multi-machine deployment.
3) **Generative AI Function Calling** effectively translates natural language into robotic functions.
4) **Metal-Gear Servos (MG90S)** provide the necessary durability for continuous industrial operation with payloads up to 500g.

B. Clinical Applications: Prosthetics & Rehabilitation

Beyond industrial automation, the core technology of Project AURA holds significant promise for the medical field, specifically in low-cost prosthetics and rehabilitation.

1) Democratizing Assitive Technology: Traditional myoelectric (EMG) prosthetics are prohibitively expensive ($> \$10,000$) and require extensive training to isolate muscle signals [26]. AURA’s vision-based control paradigm offers a ”Sensorless” alternative. A camera mounted on the user’s shoulder or glasses could track their *unaffected* hand and mirror its movements to the prosthetic hand (Mirror Therapy). This is particularly valuable for bilateral amputees or patients with hemiplegia.

2) Phantom Limb Pain Therapy: Phantom Limb Pain (PLP) affects up to 80% of amputees. Mirror Box Therapy is the standard treatment, but it requires the patient to be static. AURA enables ”Digital Mirror Therapy,” where the robotic hand moves in real-time synchronization with the patient’s intent (inferred from the contralateral limb), potentially inducing the cortical reorganization necessary to alleviate PLP [27].

3) Telerehabilitation: The IoT infrastructure allows for remote monitoring of rehabilitation progress. Therapists can prescribe specific ”Exercise Loops” (e.g., repeating a gripping motion) via the dashboard, and the system logs range-of-motion (ROM) metrics in real-time, enabling data-driven physical therapy [28].

C. Future Works

To further enhance Project AURA’s capabilities, future iterations will focus on:

- **Motion Library Persistence:** Integrating recorded sequences with a permanent cloud library for cross-site deployment.
- **Scheduling Engine:** allowing operators to schedule specific motions for specific shifts (e.g., 9 AM to 5 PM).
- **Multi-Robot Coordination:** Enabling synchronized looped motions across multiple machines for complex assembly line workflows.
- **WebSocket Communication:** Replacing HTTP polling with WebSockets to further reduce latency.

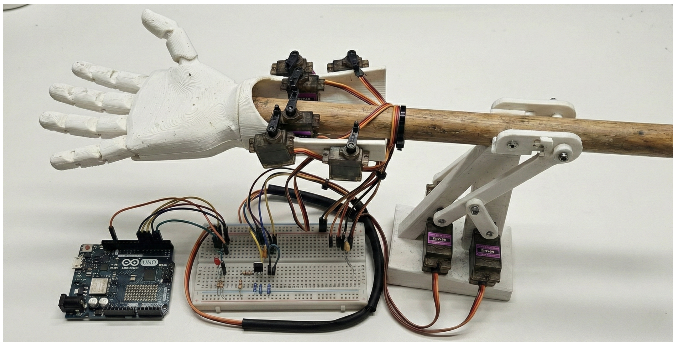

Fig. 6. Conceptual visualization of the AURA system in a clinical prosthetic setting, illustrating the Mirror Therapy data flow.

In conclusion, Project AURA demonstrates that advanced robotics can be made accessible to non-technical operators through thoughtful interface design, providing a blueprint for the democratization of industrial automation.

REFERENCES

[1] S. Fu and P. C. Bhavsar, “Robotic arm control based on Internet of Things,” in *2019 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, 2019, pp. 1–6.

[2] N. Kumar, B. Dey, C. Chetri, and A. Biswas, “Surveillance Robocar Using IoT and Blynk App,” in *Advances in Communication, Devices and Networking, LNEE*, vol. 662, Springer, 2020, pp. 423–431.

[3] H. Li *et al.*, “Tactile SoftHand-A: 3D-Printed, Tactile, Highly-underactuated, Anthropomorphic Robot Hand with an Antagonistic Tendon Mechanism,” *arXiv preprint arXiv:2406.12731*, 2024.

[4] F. Yumbla *et al.*, “An open-source 3D-printed three-fingered robotic gripper for adaptable and effective grasping,” *Biomimetics*, vol. 10, no. 1, Art. 26, 2025.

[5] H. Zhou *et al.*, “The MOTIF hand: A robotic hand for multimodal observations with thermal, inertial, and force sensors,” *arXiv preprint arXiv:2506.19201*, 2025.

[6] Arduino, “Arduino UNO R4 WiFi Datasheet,” *Arduino Documentation*, 2023. [Online].

[7] F. Zhang *et al.*, “MediaPipe Hands: On-device Real-time Hand Tracking,” *arXiv preprint arXiv:2006.10214*, 2020.

[8] OpenAI, “Function calling and other API updates,” *OpenAI Blog*, Jun. 13, 2023. [Online].

[9] Standard Bots, “Universal Robots price guide: What to expect (new and used costs),” Oct. 3, 2025. [Online].

[10] Universal Robots, “UR3e: Ultra-lightweight, compact cobot,” 2025. [Online].

[11] The Mechatronics Blog, “Top 5 Affordable Robotic Arms in 2025: Best Low-Cost Cobots,” Oct. 11, 2024. [Online].

[12] BCN3D Technologies, “BCN3D Moveo: A fully Open Source 3D printed robot arm,” Jul. 28, 2016. [Online].

[13] B. Wessling, “Wandelbots raises $84M for no-code robotics programming,” *The Robot Report*, Jan. 25, 2022. [Online].

[14] I. Bonifacic, “Wandelbots raises another $84 million to teach robots without code,” *TechCrunch*, Jan. 25, 2022. [Online].

[15] Canvas Business Model, “Who Owns Wandelbots Company?,” Jul. 12, 2025. [Online].

[16] Xpert Digital, “Wandelbots Robotics start-up and its no-code approach ’Nova’,” Jul. 23, 2025. [Online].

[17] G. García-Gil, G. del Carmen López-Armas, and J. de Jesús Navarro, Jr., “Human-Machine Interaction: A Vision-Based Approach for Controlling a Robotic Hand Through Human Hand Movements,” *Technologies*, vol. 13, no. 5, Art. 169, Apr. 2025. [Online].

[18] W. Rhee *et al.*, “Unconstrained lightweight control interface for robot-assisted minimally invasive surgery using MediaPipe framework and head-mounted display,” *Virtual Reality*, vol. 28, no. 2, May 2024.

[19] M. Wameed, A. Alkamachi, and E. Erçelebi, “Tracked Robot Control with Hand Gesture Based on MediaPipe,” *Al-Khwarizmi Eng. J.*, vol. 19, no. 3, pp. 56–71, Sep. 2023.

[20] QVIRO, “Universal Robots UR5e Cobot Reviews, Price, Use cases,” 2025. [Online].

[21] Google Developers, “Prosthesis control via Mirru App using MediaPipe hand tracking,” *Google Developers Blog*, May 26, 2021. [Online].

[22] BCN3D Technologies, “BCN3D-Moveo: Open Source 3D Printed Robotic Arm for educational purposes,” *GitHub*, Jul. 28, 2016. [Online].

[23] BCN3D Technologies, “BCN3D MOVEO - A fully OpenSource 3D printed Robot Arm,” *Thingiverse*, Jul. 28, 2016. [Online].

[24] G. Chebly *et al.*, “Strong, Accurate, and Low-Cost Robot Manipulator,” *arXiv preprint arXiv:2507.15693*, Jul. 21, 2025.

[25] A. Vaswani *et al.*, “Attention is all you need,” in *Advances in Neural Information Processing Systems*, 2017, pp. 5998–6008.

[26] F. Cordella *et al.*, “Literature review on needs of upper limb prosthesis users,” *Frontiers in Neuroscience*, vol. 10, p. 209, 2016.

[27] J. Foell, R. Bekrater-Bodmann, M. Diers, and H. Flor, “Mirror therapy for phantom limb pain: brain changes and the role of body representation,” *European Journal of Pain*, vol. 18, no. 5, pp. 729–739, 2014.

[28] A. Hernandez *et al.*, “Virtual Reality-Based Rehabilitation as a Feasible and Engaging Tool for the Management of Chronic Poststroke Upper-Extremity Function Recovery: Randomized Controlled Trial,” *JMIR Serious Games*, vol. 10, no. 3, p. e37506, 2022.