Design and Implementation of a Multifunctional Desktop Pet Robot Dog Based on Arduino Nano and ESP32-S3

Di Li*, Junkai Lin*, Siqi Hou*, Yanyan Ji** *Paper ID: 11*

Beijing Normal–Hong Kong Baptist University, Zhuhai, China

** These authors contributed contributed equally to this work.*

*** Corresponding author:* ***yyji@uic.edu.cn***

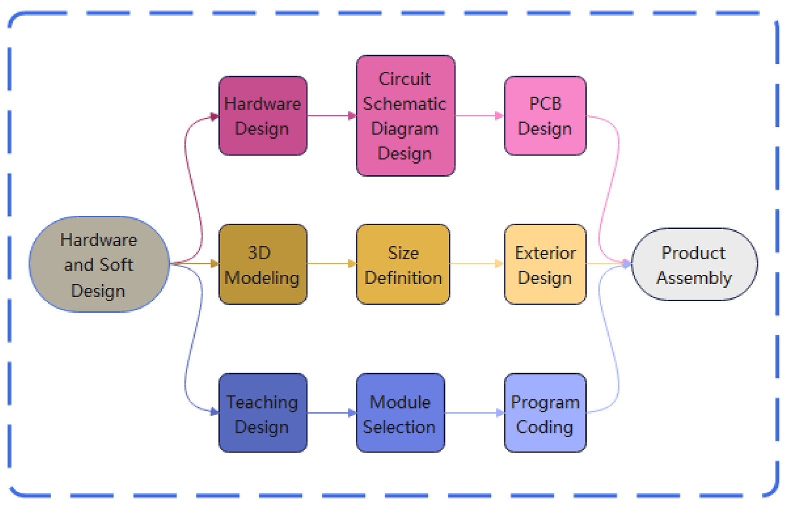
ICIRA 2025
AI & Robotics for Smart Society

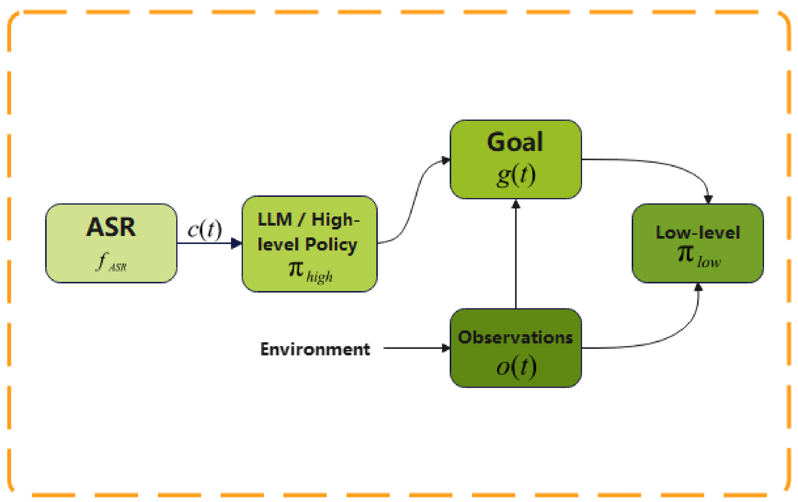
Abstract

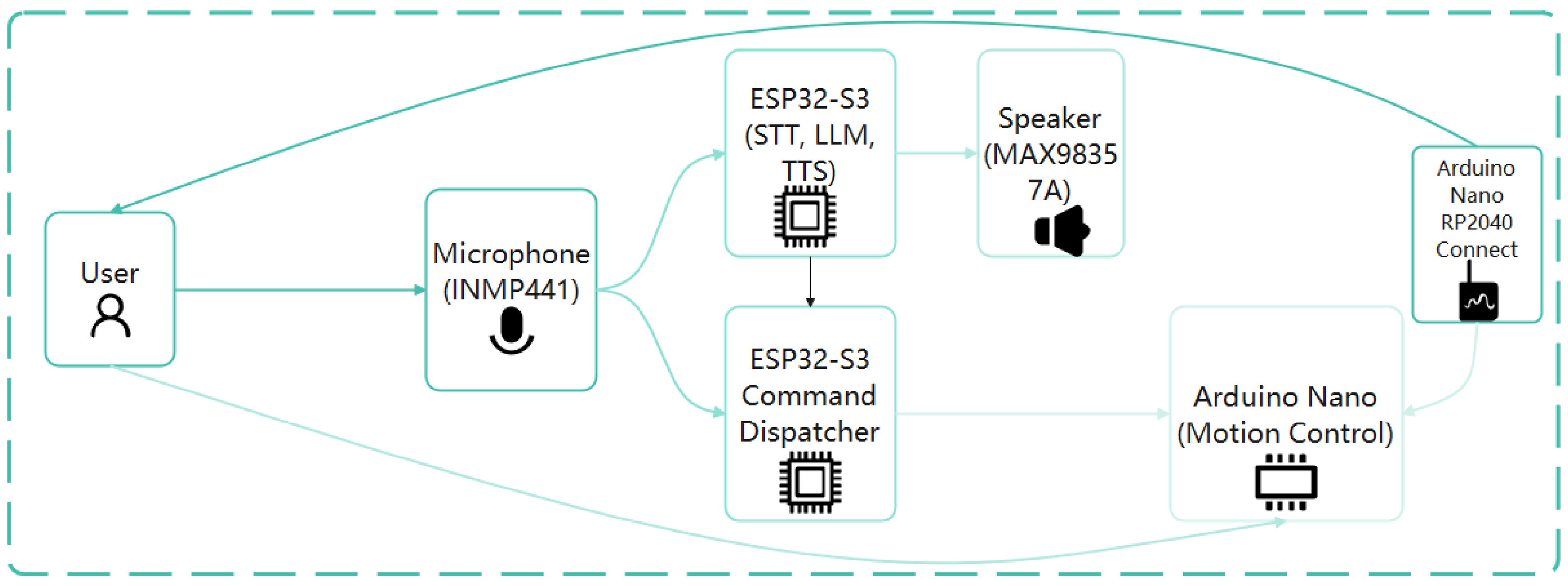
This paper designs and implements Robot Dog, a multifunctional desktop pet robot with low cost and extensibility. It runs on Arduino Nano, ESP32-S3-DevKitC-1, and Arduino Nano RP2040 Connect which support natural language interaction with large language model (LLM) and automatic speech recognition with ASR module. Furthermore, Bluetooth is used for manual control and infrared sensor is used to detect cliff and avoid falling. A novel tri-controller architecture is proposed such that the RP2040 Connect is used to implement IoT, i.e., to monitor the states of Robot Dog and send fall alerts to Arduino Cloud through IoT controller. The ESP32-S3 calls Tongyi Qwen LLM API to translate user dialogue represented by natural language to walking or greeting. The Robot Dog system with sensor fusion, hardware integration and layered interaction model: high-level goal driven by LLM and low-level servo control command at 50 Hz. Experiments confirm the system's effectiveness for AI-powered interaction on embedded platforms, with applications in education and entertainment.

Introduction

Robot Dog is a **low-cost**, **multifunctional**, and **AI-driven desktop robot dog** designed for educational use, maker projects, and human–robot interaction research. It integrates motion control, natural language understanding, and cloud-based remote monitoring.







System Architecture:

- **Arduino Nano** – Controls servo motors, OLED display, infrared and laser ranging modules.
- **ESP32-S3** – Handles intelligent interaction: ASR, NLP, TTS, and LLM (Tongyi Qwen) API communication.
- **Arduino Nano RP2040 Connect** – Enables Wi-Fi-based IoT features, IMU-based anomaly detection and Arduino Cloud integration.

Key Features:

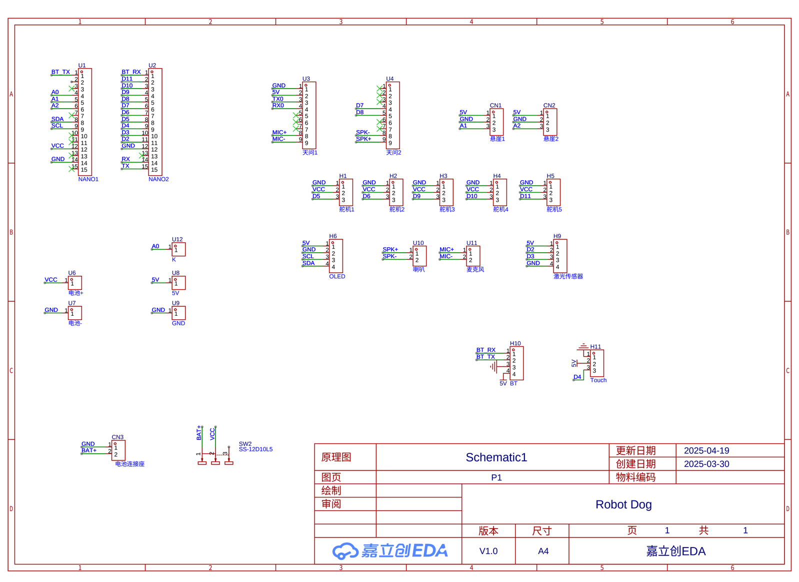
- Voice interaction using **ASR + LLM-based semantic interpretation**.
- Multimodal control: **voice commands**, **Bluetooth remote**, and **cloud-based interface**.
- Smart behavior: **hand-following** via laser sensor, **cliff-edge detection** via infrared.
- **IMU-based anomaly detection** with early-warning alert system.

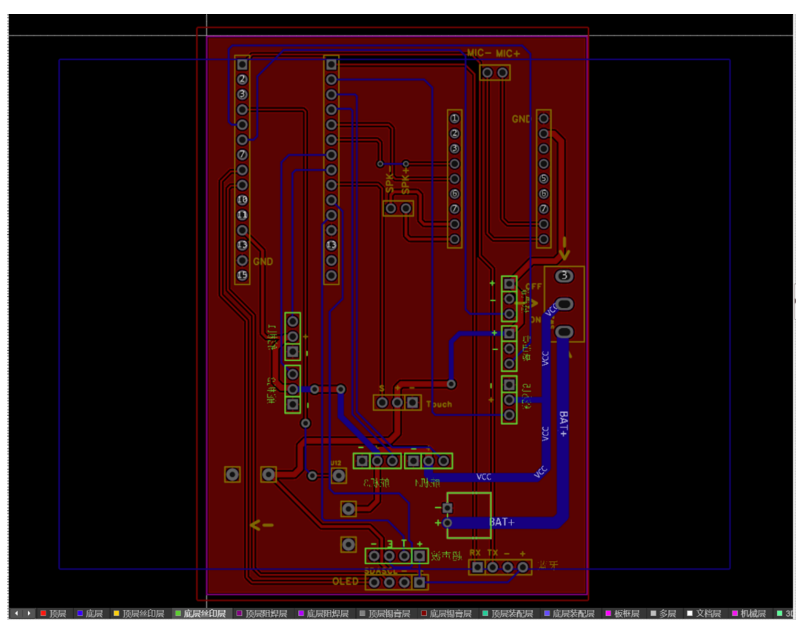
Technical Challenges:

- Multi-controller communication protocol design.
- Real-time voice interaction pipeline with behavior mapping.
- Motion sequence optimization for smooth servo control.
- Remote monitoring and alert system via Arduino Cloud.

Materials and Methods

Hardware Design



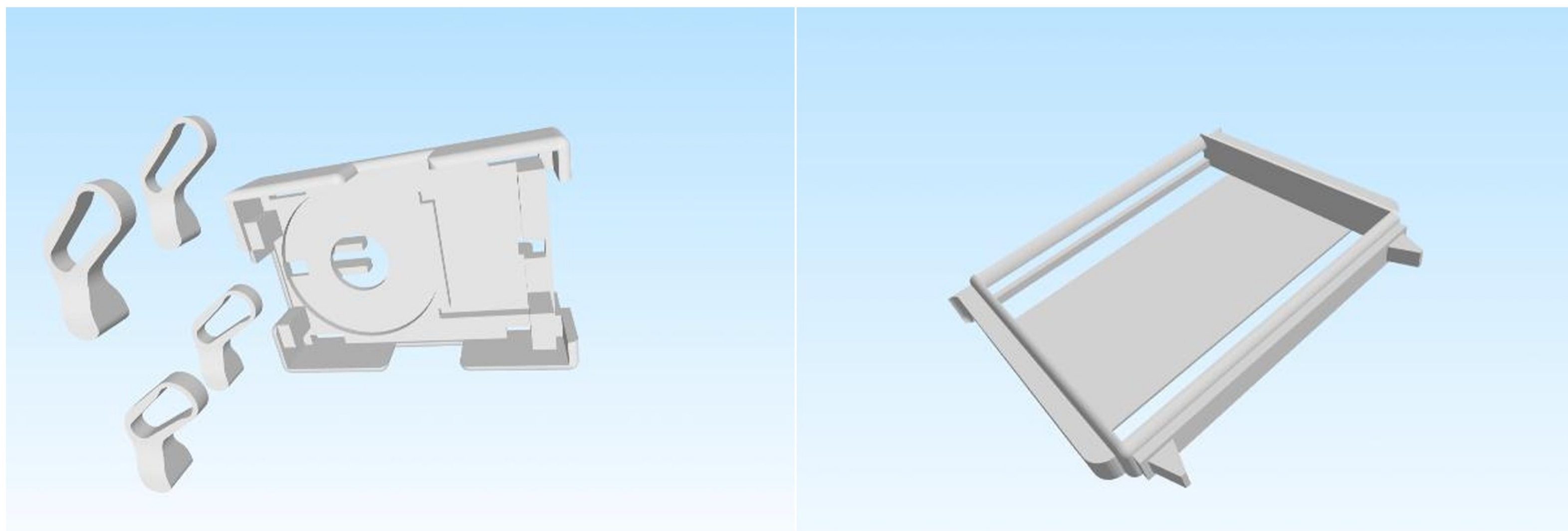


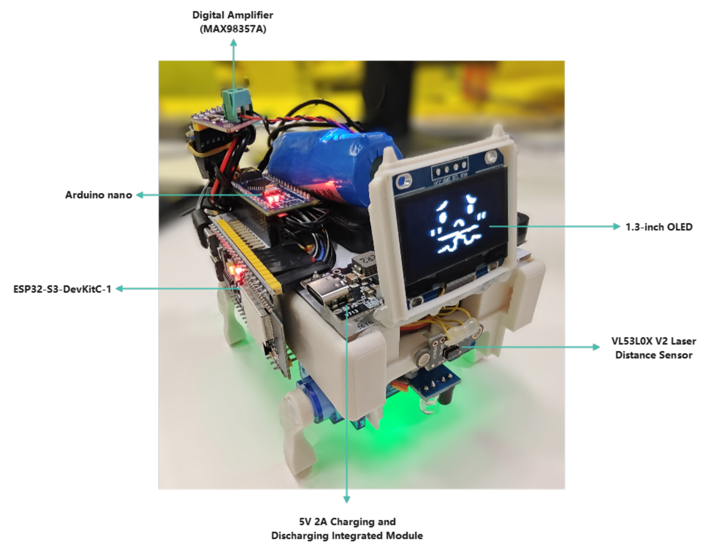
The development of Robot Dog involves a tightly integrated hardware and software architecture, combining multiple microcontrollers, sensors, and actuators to achieve intelligent behavior and robust user interaction.

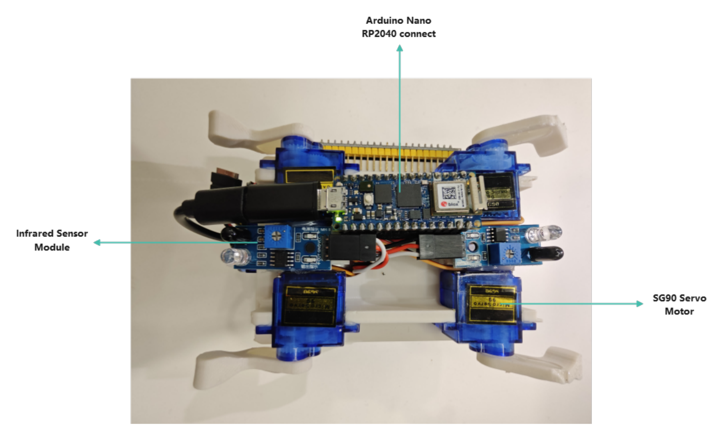
Control Units

The main motion control is handled by a Arduino Nano that handles coordination of servo actions, executing command and sensor data. The main processing and communication coordination is handled by a ESP32-S3-DevKitC-1 that communicates with the Tongyi Qwen LLM API to handle natural language understanding, command generation, and cloud coordination. An additional Arduino Nano RP2040 Connect provides the IoT connectivity to allow for remote monitoring and control of the device through the Arduino Cloud, as well as status reporting and alerting of issues via the built-in inertial measurement unit (IMU) sensor.

3d Model Design







Experiential Design & Result

Table 1. Command Recognition Accuracy (20 Trials)

System	Environment	Accuracy
Offline ASR	Quiet	80%
Offline ASR	Moderate Noise (55 dB)	65%
Online LLM Intent	Quiet	85%

Table 1 shows that the online LLM-based intent recognition outperforms offline ASR in accuracy and noise robustness.

Table 2. Sensor System Performance

Test	Conditions	Success Rate / Performance Metric
Cliff Detection	10 trials, varied speed/angle	80% Success Rate
Hand-Following	Maintain 10 cm distance	Avg. steady-state error: ±2.0 cm
Fall Detection (IMU)	10 simulated falls	70% Detection & Alert Transmission

Table 2 demonstrates that the robot's safety and interaction sensors perform reliably overall, with reasonable accuracy in cliff detection, precise hand-following, and effective fall alerting.

Table 3. Concise Comparison of Desktop Robot Architectures

Feature	Robot Dog (This Project)	XGO-Mini2	ESP-Demo
Architecture	Tri-Controller	Single-Controller	Single-Controller
AI Model	Hierarchical Closed-Loop	Preset Scripts	Linear Open-Loop
Sensor Loop	Dual-Layer (Informs AI)	Low-Level (For Stability)	None in AI Loop
Core Advantage	Adaptive Interaction	Advanced Mobility	LLM Proof-of-Concept

Table 3 highlights the Robot Dog's superior adaptability and autonomy enabled by its tri-controller architecture and sensor-informed AI, setting it apart from XGO-Mini2 and ESP-Demo.

Conclusion

The Robot Dog project demonstrates the feasibility of creating a compact, intelligent desktop pet robot through modular hardware-software co-design. By integrating technologies such as servo actuation, speech recognition, and IoT monitoring, it validates the use of embedded AI in consumer robotics. The system's modular architecture, covering areas like motion control and speech processing, simplifies development and enhances scalability. Iterative prototyping and testing improved hardware stability and software robustness, with safety features like infrared cliff-edge detection. A notable achievement is the integration of large language models for natural language interaction, enabling meaningful dialogues that translate into behavioral responses. The project exemplifies an interdisciplinary approach, combining robotics, AI, and user-centered design. Future work will focus on adding behavioral learning, customizable interfaces, and facial recognition. Overall, Robot Dog provides a reference for designing low-cost, extensible, and intelligent personal robotics systems.