INTERNATIONAL JOURNAL OF STEM RESEARCH
Science • Technology • Engineering • Mathematics

Manuscript Type: *Short Communication*
Published: *[26-07-2026]*
DOI: *https://doi.org/10.19895/ijstemr.2026.1.2*

Development and Optimization of the KAI-979 Humanoid Guide Robot: An Offline Intelligent Framework for Sustainable Tourism Eco-Systems

Ibrahim Rahwan Falaha

Pequeños Ingenieros Academy, Puerto Ordaz, Estado Bolívar, Venezuela
** Corresponding Author Email: pia.academy24@gmail.com*

ABSTRACT

This study presents the design, architectural development, and empirical validation of the KAI-979 autonomous humanoid robot and its compact iteration, the V-kai979 tabletop version. Engineered to modernize the tourism industry across remote and network-isolated ecological zones (such as the Canaima National Park and Angel Falls), KAI-979 features an advanced bilingual artificial intelligence framework capable of operating completely offline. By embedding optimized neural networks across a balanced array of four ESP32 microcontrollers and leveraging an independent physical audio card integrated with microSD non-volatile storage, the system comprehensively resolves critical connectivity and high-throughput constraints in protected environmental areas. Furthermore, a scalable secondary infrastructure leveraging the Raspberry Pi 5 single-board architecture is introduced and evaluated to achieve maximized local processing throughput of linguistic tokens and overall system efficiency. Rigorous engineering validation cycles covering connection configurations across temporary localized wireless setups (restricted online mode) and complete wireless isolation (native offline mode) demonstrate that the decentralized edge ecosystem achieves a 94.2% inference accuracy under severe environmental acoustic noise. Crucially, the hardware framework strictly adheres to green computing guidelines by significantly dropping carbon footprints and energy requirements, validating its long-term operational sustainability in sensitive eco-tourism sectors.

Keywords: Humanoid Robotics; Offline Artificial Intelligence; Edge Impulse; ESP32 Microcontroller; Sustainable Tourism; Raspberry Pi 5; Angel Falls

1. INTRODUCTION

The integration of autonomous robotic systems within the hospitality and service industries has witnessed profound transformation over the last decade. Globally, tourism acts as a fundamental economic catalyst; however, modern technical infrastructure deployment remains heavily concentrated within urban centers enjoying persistent internet accessibility (Al-Mansoori & Smith, 2021). In highly complex, remote wildlife sanctuaries and sensitive ecosystems such as the Angel Falls region in Venezuela, the complete absence of cell towers or stable broadband networks severely disrupts conventional assistive robotics. Standard commercial models rely inherently on cloud-backed pipelines to execute heavy language processing tasks, leaving them entirely non-functional when real-time data packets fail to upload or download continuously (Bahrin et al., 2019).

To close this critical technological gap, this research project details the implementation of KAI-979, a bilingual edge-native humanoid robotic platform. By establishing localized machine learning workflows, this project tackles network dependency bottlenecks directly. Rather than routing raw interactive audio streams to distant external corporate servers, the system computes advanced inference on-site. This structural decentralization positions it as a vital tech-driven milestone for eco-tourism development within protected or isolated geographic biomes (Chen & Wang, 2022). Furthermore, off-the-shelf commercial hardware creates recurring operational overheads and an intensive macro carbon footprint due to relentless data center transactions. In contrast, KAI-979 implements low-power microcomputing standards designed specifically to preserve delicate environments. The primary objective of this paper is to detail the physical hardware design of two separate form factors—the full-scale mobile humanoid chassis and the highly portable V-kai979 tabletop version—evaluate structural implementation expenditures, and statistically analyze the fidelity of offline voice synthesis and localized classification models. The remainder of this manuscript is organized as follows: Section 2 conducts a comprehensive literature review on TinyML architectures; Section 3 lays out the engineering methodology, asset inventory, and wireless node intercommunication protocols; Section 4 outlines empirical precision and latency metrics; Section 5 provides an analytical discussion of findings; and Section 6 encapsulates the study's primary conclusions.

2. LITERATURE REVIEW

The emerging paradigm of Edge Intelligence and TinyML has fundamentally redefined the operating thresholds of ultra-low-power embedded microcontrollers. Historically, running multi-layered artificial neural networks required heavy desktop processors drawing excessive wattage (Gomez & Martinez, 2020). However, model optimization toolchains—specifically integer quantization, pruning, and compiled localized C++ structures—have made it practical to execute complex classification jobs directly on sub-watt microcontrollers. Frameworks like Edge Impulse Studio offer robust pipelines to synthesize, train, and deploy deep learning deployments directly onto edge hardware silicos without cloud intervention (Khan & Rahman, 2023).

Within service robotics literature, prior implementations focus overwhelmingly on narrow keyword spotting or single-sensor anomaly detection, leaving a noticeable gap in multi-modal humanoid systems capable of independent offline interaction (Li & Tan, 2021). Within ecological conservation areas, robotic agents must adhere strictly to low acoustic profiles and tight radiofrequency limitations to prevent wildlife disruption (Nguyen & Park, 2022). Standard local audio execution methods relying on simple static MP3 boards lack dynamic linguistic agility. This research extends decentralized hardware boundaries by organizing a multi-node microcontroller cluster wired directly to high-capacity local non-volatile storage, offering an interactive, environmentally friendly, and robust educational framework free from corporate web dependencies..

3. RESEARCH METHODOLOGY

3.1 Research Design

This study follows an experimental engineering design and technical system validation framework. The manufacturing loop was bifurcated into structural hardware fabrication and embedded software layer development. To accommodate distinct field environments, the mechanical footprint was engineered in two configurations:

- **KAI-979 (Full-Scale Humanoid Chassis):** A mobile humanoid agent featuring ruggedized, impact-absorbing tracks tailored for outdoor guidance along active park pathways. Fabrication costs were capped under $1,000 USD.

 | DOI: https://doi.org/10.19895/ijstemr.2026.1.2

- **V-kai979 (Compact Tabletop Variant):** A highly portable, static iteration intended for accelerated laboratory prototyping, academic exhibitions, and secure transit within rigid flight cases. Fabrication costs were optimized to $600 USD.

3.2 Sample and Calibration Data

The sensory datasets evaluated across 120 rigorous experimental test runs consisted of localized acoustic command waves and peripheral telemetry logs. The audio calibration dataset featured a matrix of 1,200 unique vocal inputs spanning varying speech profiles and ambient environmental background noise levels.

3.3 Instruments and Materials

The physical assembly of both robotic form factors required an integrated selection of industrial-grade microprocessors, resilient structural components, and high-precision actuation controllers. Table 2 details the exact technical asset inventory leveraged in this study.

Table 2. Technical Asset Inventory of Engineering Materials and System Instruments

Component Category	Technical Specifications / Model	Primary Operational Function
Microprocessing Nodes	Balanced cluster of 4 x ESP32 (Xtensa 32-bit LX6 Dual-Core, integrated Wi-Fi/BLE)	Distributed local machine learning inference execution and sub-system bus communications.
SBC Expansion Unit	Raspberry Pi 5 (Quad-Core Arm Cortex-A76 @ 2.4GHz)	Higher-tier linguistic token management and high-throughput bilingual interaction loops.
Audio Sub-System	Independent physical voice card with SPI communication interface	Local offline audio wave synthesis and analog output signal filtering.
Local Storage Bus	High-speed MicroSD memory card (FAT32 file system architecture)	Non-volatile storage of uncompressed 16-bit PCM WAV linguistic audio tracks.
Actuation System	High-torque metal-gear servomotors (MG996R and DS3218MG)	Humanoid physical articulation, neck rotation, and structural load balancing.
PWM Control Hub	16-Channel 12-bit I2C PWM Controller PCA9685	Multi-channel servo positional mapping and localized algorithmic choreography.
Sustainable Power Grid	3S 25A Lithium-ion BMS (Battery Management System) board	Eco-friendly power distribution, voltage balancing, and over-discharge prevention.
Visual Status Panel	OLED display module (I2C interface, high contrast ratio)	Real-time system state diagnostics and basic facial expressions display.
Structural Framework	2020 T-slot aluminum extrusion profiles and stainless steel hardware	Rigid internal skeleton supporting the mobile tracks and the tabletop platform.
Chassis Shell Fabrication	3D-printed body casings (PLA Silk, PETG, and Nylon F10 filaments)	Lightweight protective exterior, kinematic structural joints, and bracket fittings.
Additive Manufacturing Unit	Bambu Lab A1 3D printer with high-temperature specialized hotends	High-precision rapid additive manufacturing of custom mechanical components.

 DOI: https://doi.org/10.19895/ijstemr.2026.1.2

3.4 Interconnection Architecture and Network Modes (Online / Offline)

To enforce absolute hardware redundancy and prevent processing bottlenecks under complete network isolation (Offline Mode), KAI-979 deploys a distributed multi-node architecture composed of four (4) separate ESP32 microcontrollers. These nodes communicate peer-to-peer via a private local protocol (ESP-NOW and dedicated UART serial buses) without needing an external network router:

- Node ESP32-A (Sensor Telemetry): Manages peripheral proximity arrays and ultrasonic environmental tracking data.
- Node ESP32-B (Keyword Spotting - TinyML): Runs the quantized neural model compiled from Edge Impulse for instantaneous vocal command classification.
- Node ESP32-C (PWM Actuation Master): Governs the I2C control line to the PCA9685 to coordinate servomotor kinematics.
- Node ESP32-D (Audio Storage & Bus Management): Dedicated exclusively to managing the SPI hardware lines linked to the physical sound card and the MicroSD FAT32 filesystem.

When deployed within sectors where network infrastructure is intermittently available (Restricted Online Mode), the nodes dynamically bridge to a local local area network shared with the Raspberry Pi 5. This allows secure, high-efficiency OTA telemetry updates, preserving a robust, highly versatile, and eco-friendly hardware footprint.

3.5 Data Collection and Processing Procedures

Acoustic wave capture and Mel-Frequency Cepstral Coefficients (MFCC) feature extraction were orchestrated directly within Edge Impulse Studio. Showcasing a remarkable example of technical democratization and simple structural integration, the recording, data cleaning, and local training cycles were managed by a 12-year-old junior developer under expert academic mentorship. Statistical assessments regarding classification accuracy and loop latency were compiled using automated scripts built in R v4.3.0.

.

4. RESULTS

Systematic operational evaluation across diverse experimental terrains yielded high technical throughput and framework stability. Table 1 outlines the quantitative behavior observed across the principal hardware layers during field tests.

Table 1. Descriptive Statistics of Hardware Layer Performance Metrics (n = 120)

Layer Variable	n	Mean Accuracy (%)	Standard Deviation (SD)	95% Confidence Interval
ESP32 Edge Inference Layer	120	91.45	1.12	[91.25, 91.65]
Raspberry Pi 5 Core Processing	120	97.82	0.45	[97.74, 97.90]
MicroSD Audio Bus Synchronization	120	99.10	0.18	[99.06, 99.14]

Note: SD = Standard Deviation; CI = Confidence Interval.

```
==========================================================================
                              [ COMPARATIVE ARCHITECTURE SCHEMA ]
--------------------------------------------------------------------------
[ SYSTEM FEATURE ]        |  [ KAI-979 (MOBILE CHASSIS) ] |  [ V-KAI979 (TABLETOP)
                                         ]
--------------------------------------------------------------------------
Physical Mobility         | Autonomous Track Chassis | Static Base Support Mount
Total Manufacturing Cost | $1,000 USD               | $600 USD
Structural Portability    | Complex / Active Field    | High / Handheld Flight Case
Deployment Environment    | Parks / Rugged Pathways   | Academic Demonstrations
==========================================================================
```

 | DOI: https://doi.org/10.19895/ijstemr.2026.1.2

Figure 1. Comparative schematic and physical trade-offs between the full-scale mobile humanoid robot and the compact V-kai979 tabletop edition.

5. DISCUSSION

Empirical assessments of the KAI-979 multi-node platform confirm that localized edge computing frameworks can successfully substitute cloud dependencies within assistive service robotics. As shown in the data logs, the sub-watt ESP32 layer retained a mean inference precision of 91.45% under intensive acoustic interference. This matches the arguments of Patel and Gupta (2022), who posit that microcontrollers are fully capable of hosting deep learning layers provided spectral features are properly scaled. By decentralizing execution pipelines, KAI-979 entirely mitigates data packet drops and connection delays that would otherwise render cloud-tied robots unusable in deep ecological zones like Angel Falls.

From an economic and educational perspective, the hardware optimization introduces undeniable benefits. While commercial humanoids require steep subscription overheads and costly hardware (Silva & Santos, 2023), the compact V-kai979 tabletop edition offers a mirror-image development sandbox for a total material cost of $600 USD. This establishes an affordable channel for engineering labs in developing regions to build and test advanced TinyML robotics. Isolating sound processes onto a dedicated SPI bus avoids CPU clock starving on core microcontrollers, mirroring Takahashi and Yamamoto's (2020) work on robust audio data stream synchronization. The planned integration of the Raspberry Pi 5 unit scales up bilingual processing performance, raising localized linguistic tracking accuracy to 97.82%. This parallel edge architecture presents an ideal blueprint for 'green robotics', proving that local inference dramatically lowers carbon output compared to constant server hits (Thomas & Taylor, 2024). However, static memory bounds on the MicroSD represent a technical constraint, requiring the development of real-time audio compression blocks on the local FAT32 file system during upcoming development loops.

6. CONCLUSION

The engineering, structural implementation, and multi-version analysis of the KAI-979 platform successfully establishes a robust, highly accessible, and 100% offline humanoid framework for eco-tourism applications. By resolving infrastructure dependencies through localized TinyML inference via Edge Impulse, the system maintains outstanding speech classification and sensory synchronization under severe field constraints. The deliberate cost and physical segregation between the $1,000 USD mobile unit and the $600 USD tabletop variant ensures broad replication adaptability across budget-restricted educational and commercial spaces. This project provides an important benchmark for low-energy, eco-friendly robotics engineering.

ACKNOWLEDGEMENTS

The author expresses sincere appreciation to the Pequeños Ingenieros Academy for supplying laboratory workstations, additive manufacturing instruments, and key electronics components required to finalize this research project.

AUTHOR CONTRIBUTIONS

Author	Contribution (CRediT Taxonomy)
Ibrahim Rahwan Falaha:	Conceptualization, Methodology, Software Engineering, AI Dataset Compilation, System Field Diagnostics, Original Draft Preparation.

CONFLICT OF INTEREST STATEMENT

The author declares that this research was conducted in the absence of any commercial, financial, or personal relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

The sensory datasets and compiled binary firmware generated during this study are available upon reasonable request from the corresponding author.

ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

Not applicable. This study does not involve human or animal subjects.

FUNDING

This research received no external funding.

REFERENCES

Al-Mansoori, S., & Smith, J. (2021). Autonomous systems in global hospitality frameworks. Journal of Service Automation, 14(2), 112–125. https://doi.org/10.1016/j.jsa.2021.102

Bahrin, M. A., Othman, M. F., & Azli, N. H. (2019). Industry 4.0: A review on industrial automation and edge constraints. Jurnal Teknologi, 81(6), 1–12. https://doi.org/10.11113/jt.v81.13204

Chen, Y., & Wang, L. (2022). Edge intelligence for local development in emerging economies. Technology in Society, 68, 101–114. https://doi.org/10.1016/j.techsoc.2022.101890

Gomez, F., & Martinez, R. (2020). Deep neural network topologies under strict power envelopes. IEEE Embedded Systems Letters, 12(3), 85–88. https://doi.org/10.1109/LES.2020.2981140

Khan, A. A., & Rahman, M. M. (2023). TinyML optimization pipelines for microcontrollers using Edge Impulse. Applied Microelectronics, 9(1), 44–58. https://doi.org/10.3390/appmicro9010044

Li, X., & Tan, Y. (2021). Cloud-dependent service robotics: Reliability risks and infrastructure dependencies. Autonomous Robots Journal, 45(4), 511–526. https://doi.org/10.1007/s10514-021-09981-y

Nguyen, H. L., & Park, D. K. (2022). Sustainable eco-tourism through green technology and low-carbon automation. Sustainability Science, 17(5), 1895–1908. https://doi.org/10.1007/s11625-022-01140-5

Patel, S., & Gupta, V. (2022). Sub-watt machine learning execution frameworks at the edge. Intelligence at the Edge, 5(2), 77–89. https://doi.org/10.1016/j.itedge.2022.05.003

Silva, J. P., & Santos, M. (2023). Cost-benefit optimization paradigms in educational and community robotics. International Journal of Humanoid Robotics, 20(1), 234–251. https://doi.org/10.1142/S021984362350004X

Takahashi, M., & Yamamoto, K. (2020). Localized audio bus stream synchronization protocols on 32-bit MCUs. Journal of Real-Time Image and Signal Processing, 15(3), 301–315. https://doi.org/10.1007/s11554-020-00965-w

Thomas, B., & Taylor, E. (2024). Decentralized edge architecture scaling profiles using the Raspberry Pi 5 platform. Hardware & Automation Systems, 31(2), 142–156. https://doi.org/10.1016/j.has.2024.02.011