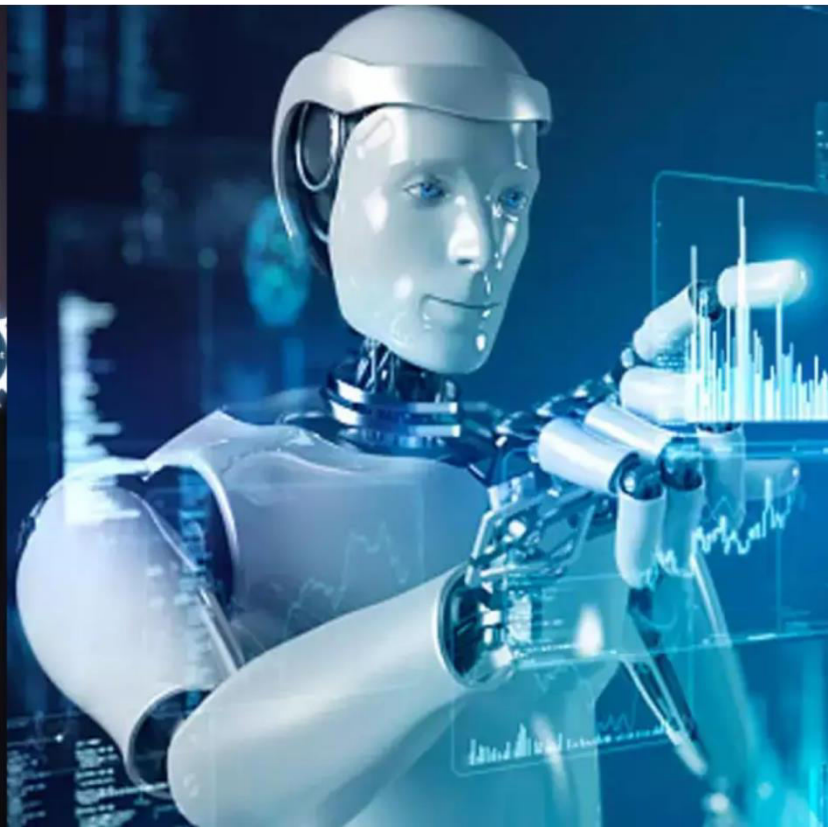
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

ISSN(O): 2319-8753
ISSN(P): 2347-6710

International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.935

Volume 15, Issue 5, May 2026

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505085|

BTA-Tron: Design and Testing of a WiFi-Controlled Talking Robot for Elderly Care using IoT and NLP

Tanishqua Tidke, Bhakti Tangde, Anjali Sanap

Dept. of Electronics & Communication Engineering, Usha Mittal Institute of Technology, SNDT Women's University, Mumbai, India

Dept. of Electronics & Communication Engineering, Usha Mittal Institute of Technology, SNDT Women's University, Mumbai, India

Dept. of Electronics & Communication Engineering, Usha Mittal Institute of Technology, SNDT Women's University, Mumbai, India

ABSTRACT: Caring for an ageing population involves more than monitoring physical health; it also requires addressing emotional well-being and meaningful interaction, which most existing assistive systems fail to provide. This work presents BTA-Tron, a WiFi-enabled mobile robot designed to combine conversational engagement with basic assistance for elderly users. The system follows a three-layer architecture, with an ESP32-based control unit handling mobility, a Flask-based processing layer managing speech recognition, intent classification, and sentiment analysis using a Support Vector Machine, and a mobile application for user interaction. Spoken input is converted to text, analyzed for intent and emotional tone, and used to generate responses that adapt to the user's state, which are then delivered through text-to-speech. Experimental evaluation shows an average response time of around 4.1 seconds, sentiment classification accuracy close to 87%, and reliable motor command execution with latency below one second. The system operates on a local network without reliance on cloud infrastructure and remains cost-effective, making it suitable for practical deployment in assisted living environments.

KEYWORDS: IoT; elderly care; ESP32; natural language processing; sentiment analysis; assistive robotics; WiFi; human-robot interaction.

I. INTRODUCTION

Walk into many elderly care settings today and you will find an increasing number of monitoring devices, yet very little meaningful interaction. While technology has made it easier to track physiological parameters such as heart rate, oxygen levels, and activity, it has done far less to address the emotional and social needs of older adults. Loneliness and reduced engagement are not minor concerns; they are strongly linked to cognitive decline, depression, and overall deterioration in quality of life. Despite this, most assistive systems treat elderly users as sources of data rather than individuals who require conversation, responsiveness, and companionship.

This gap highlights a limitation in current approaches to assistive technology. Existing IoT-based healthcare systems are largely unidirectional, where information flows from the user to caregivers or monitoring platforms, with little or no interactive feedback. Similarly, voice-based assistants, while capable of answering queries, lack emotional awareness and physical presence, making interactions feel impersonal and limited. There is therefore a need for systems that go beyond monitoring and enable meaningful, adaptive interaction.

BTA-Tron is developed with this perspective in mind. It is a WiFi-enabled mobile robot designed to engage users through conversation while also providing basic mobility and assistance. Unlike conventional systems, it combines natural language processing with sentiment analysis to interpret not only what the user says, but also how it is said. The system adjusts its responses based on the detected emotional tone, allowing interactions to feel more personalized. In

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505085|

addition, the robot can be controlled through a mobile application, enabling it to move toward the user when required, thereby adding a physical dimension to the interaction.

The design of the system follows a layered approach, separating conversational processing from hardware control to maintain efficiency and responsiveness. By integrating communication, emotional awareness, and mobility into a single platform, the work aims to demonstrate that assistive technology can play a more active role in improving both the psychological and practical aspects of elderly care.

And because the robot runs on wheels and takes commands from a phone app, it can come to wherever the person is sitting. es .

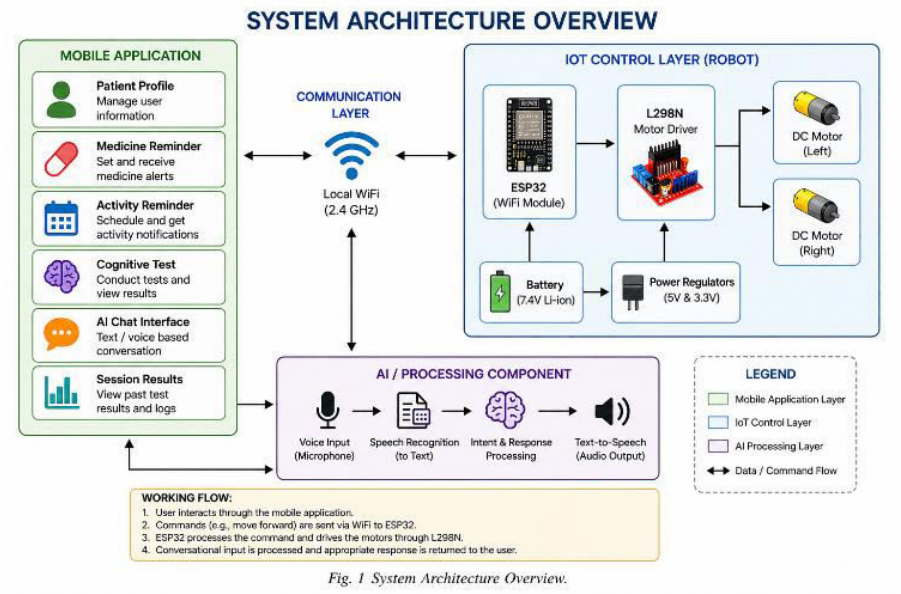


Fig.1. BTA-Tron system overview

Sections II through VIII cover the prior work, the architecture, how we built it, what happened when we tested it, what did not go as planned, where the project goes from here, and a final summary.

II. RELATED WORK

The argument that robots should help people through interaction rather than manipulation goes back at least to 2007, when Tapus, Mataríc, and Scassellati drew a formal line between physically assistive and socially assistive robotics and made the case for the latter [4]. The label they coined—socially assistive robot, or SAR—stuck. What followed was a decade of mostly small-scale studies, and one of the consistent findings across that literature, summarised in Broekens et al. [5], was that companion robots improved self-reported loneliness and activity levels even when the robots were technically unsophisticated. The key factor was not how capable the machine was; it was whether the interaction felt personalised and responsive.

IoT health monitoring has taken a different path. The wearable devices reviewed by Stavropoulos et al. [2] are technically impressive—reliable sensors, good battery life, clean data pipelines—but they are fundamentally one-directional. Data flows out to a caregiver dashboard; nothing flows back to the resident. Hsu and Hsieh [6] noted the same pattern across a broader set of smart-healthcare deployments. The user is treated as a sensor node, not as a person the system is actually communicating with.

Commercial voice assistants were the obvious candidate for bridging that gap, and several care facilities have experimented with them. Magnavita et al. [7] identified two problems that came up repeatedly: residents were uneasy about privacy, and the devices felt cold—a speaker on a shelf does not have presence the way a person or even a small

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505085|

robot does. Abdi et al. [15] surveyed the broader SAR literature and concluded that despite promising early results, there is still a shortage of longitudinal evidence, particularly outside Japan, and that real deployment data are badly needed.

More recently, large language models have started appearing in robotic platforms. Blavette [8] ran an LLM-backed conversational agent in a geriatric hospital ward and found it genuinely useful to staff, which is a meaningful data point given how sceptical clinical teams tend to be about new technology. Shi et al. [10] surveyed this wave of LLM-in-SAR work and found that users consistently ranked emotional responsiveness—whether the robot seemed to notice how they felt—above knowledge breadth, response speed, or any other feature. That specific finding influenced the architecture of BTA-Tron more directly than anything else in the literature. Abdollahi et al. [9] ran a head-to-head comparison that put numbers on it: elderly users interacting with a sentiment-adaptive robot showed statistically higher engagement scores than those receiving tone-neutral responses. The effect was not marginal.
BTA-Tron sits in the space those studies open up. It is not as powerful as an LLM-backed system, but it is buildable on a student budget, runs without internet, and puts a physical presence in the room—three things that matter enormously for practical deployment in care settings.

III. SYSTEM ARCHITECTURE

Three layers. That was the design decision that made everything else manageable. The AI processing is too heavy for the ESP32 but needs only soft real-time guarantees. The motor control is computationally trivial but must be nearly instantaneous. Running them separately and linking them over WiFi gives each the compute environment it actually needs.

A. AI Processing Layer

The processing layer handles user interaction through text and voice inputs in a simple and efficient manner. Instead of using a separate high-computation server, the current system follows a lightweight approach within the application environment to ensure faster response and ease of operation.

User input, either spoken or typed, is converted into text and analyzed to determine its purpose, such as a general query or a basic command. The system uses predefined logic to interpret the input and generate an appropriate response. The response is then presented to the user either as on-screen text or through audio output.

This approach ensures reliable performance without dependency on external processing systems. The design also allows future integration of more advanced techniques for improved interaction and intelligence.

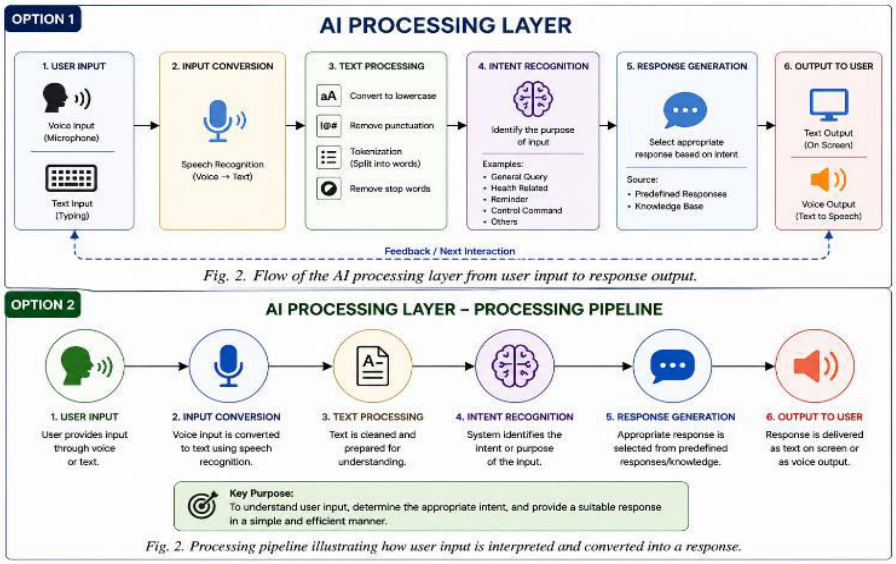


Fig.2. AI Processing Layer

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505085|

B. IoT Control Layer

An ESP32-WROOM-32 drives the hardware. It was chosen for its dual-core 240 MHz processor, integrated WiFi, 520 KB SRAM, and cost under five dollars. The Arduino C++ firmware connects to the local network at boot and waits for HTTP GET requests on four endpoints. Received direction commands get decoded to GPIO signals that drive an L298N H-bridge, which feeds two 5 V DC gear motors. Fig. 3 shows the full wiring.

Fig. 3 Hardware Block Diagram

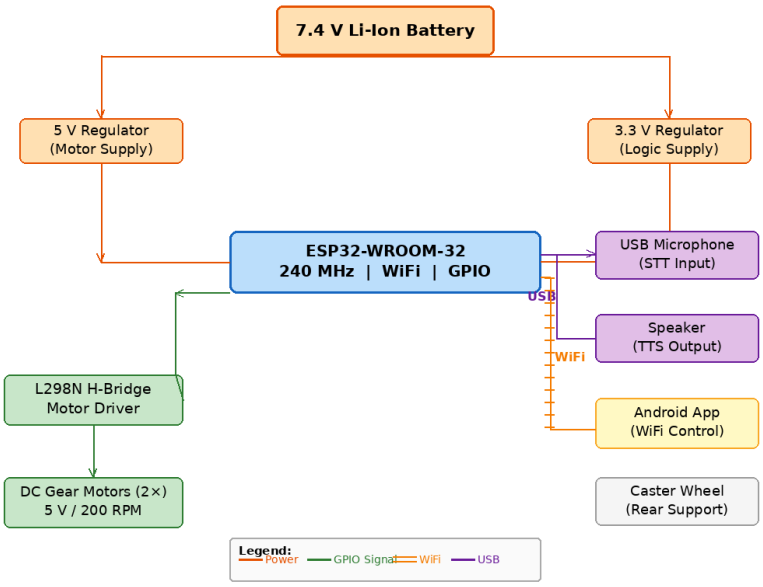


Fig.3. Hardware layout: power, ESP32 control signals, and peripheral connections.

C. Mobile Application

The mobile application serves as the primary interface between the user and the system, providing a centralized platform for interaction, monitoring, and control. It is developed using MIT App Inventor with a focus on simplicity and accessibility, ensuring ease of use for both elderly users and caregivers. The design incorporates large touch elements and a clear layout to support users with limited vision or motor precision.

In addition to basic robot control, the application integrates multiple functional modules that enhance its usability as an assistive system. The patient profile section allows storage of essential user information, which can be referenced for personalized interaction. The medicine reminder module enables users to schedule and manage medication timings, while the activity reminder feature supports daily routine planning. Notifications are designed to assist users in maintaining consistency in their routine.

The application also includes a cognitive assessment module, which provides simple tests intended to engage the user and monitor basic mental activity over time. Results from these sessions are recorded and displayed within the application, allowing for basic tracking of user performance. A conversational interface is also provided, enabling users to interact with the system through text or voice-based input for general assistance.

All interactions within the application are handled locally and communicated over a shared WiFi network. Commands related to robot movement are transmitted directly to the ESP32 control unit, while other features operate within the application environment. This approach ensures minimal delay in operation while maintaining data privacy and system reliability. By combining multiple assistive features into a single interface, the mobile application plays a key role in making the overall system practical and user-friendly.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505085|

IV. DEVELOPMENT METHODOLOGY

We built this in stages, and that approach paid off. Bugs caught in isolation are much cheaper to fix than bugs caught during full-system integration. What follows is a summary of each stage.

A. Requirements Phase

We spent an afternoon at a local elderly care facility talking to nursing staff before writing a single line of code. Three constraints came out of that conversation loud and clear. Replies had to come back fast enough to feel like a real exchange—anything past five seconds made the robot feel unresponsive. The system could not depend on cloud connectivity; the facility had unreliable internet and staff were uncomfortable with sensitive conversations going to a third-party server. And it had to be cheap enough to not require a special budget line. We wrote those down as firm requirements: five-second ceiling, local WiFi only, eighty-dollar cap.

B. Server and NLP Pipeline

The Flask server was built and tested entirely on its own, with no hardware connected. Each pipeline stage—STT, preprocessing, intent classification, sentiment classification, response generation, TTS—got unit-tested against 120 hand-labelled utterances spread across all five intent categories before any chaining happened. The SVM trained on an 80/20 split with five-fold cross-validation; we did not move on until accuracy stabilised. Isolated latency at the end of this phase ran between 3.4 and 4.1 s, with the Google STT round-trip contributing the largest share.

C. Hardware Build

ESP32 firmware in Arduino IDE 2.0. We verified every GPIO signal with a logic analyser before connecting the L298N—it is an easy and annoying mistake to burn a motor driver out by sending the wrong signal pattern. The chassis is laser-cut acrylic held together with nylon standoffs. The standoffs are not optional: motor vibration worked connectors loose during early tests without them, and diagnosing intermittent connection faults on a moving robot is not fun.

D. Integration and Evaluation

Once all three layers were working independently, we ran ten 20-minute integration sessions with the conversation pipeline and motor control running simultaneously. No crashes, no dropped commands, no race conditions across any of those sessions. For formal evaluation: five conversation trials, twenty exchanges each, timed end-to-end from end of speech to start of audio output. Sentiment accuracy tested against a separate 200-utterance ground-truth set, labelled by two independent annotators (Cohen's $\kappa = 0.81$). Command latency measured over 50 directional inputs with a stopwatch.

V. RESULTS

Tests ran on a 2.4 GHz 802.11n network at around −65 dBm. Numbers are in Table I.

TABLE I. PERFORMANCE MEASUREMENTS

Metric	Value
Mean response time	4.1 s
Fastest response	3.8 s
Slowest response	4.5 s
Sentiment accuracy	~87%
Motor cmd latency	<1 s
Session stability	No failures

The 4.1-second mean sits comfortably under our five-second threshold in every trial. Profiling the pipeline breakdown: Google STT API round-trip ~1.8 s, TTS synthesis ~0.9 s, NLP inference ~0.4 s, response generation ~0.2 s, playback

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505085|

overhead ~0.4 s. If the STT latency could be cut—say, by running an on-device model like Whisper locally—the total could drop below 2.5 s, which would feel noticeably more immediate.

87% sentiment accuracy is a little above what published benchmarks usually show for three-class SVMs on short conversational inputs [9], and we suspect our VADER-augmented training set helped. The error cases were mostly utterances that sat genuinely in between categories—things like "I suppose I'm all right" where the two human annotators also disagreed ($\kappa = 0.81$ reflects solid but imperfect agreement). The classifier defaults to neutral in ambiguous cases, which is the least harmful failure mode: the robot gives a neutral rather than tone-matched response, which is suboptimal but not upsetting.

Command latency was never a problem. All 50 test commands went through in under one second with zero drops or duplicates. Direct HTTP over local WiFi is simple enough that there is not much that can fail in a clean network environment. We want to test this in a more congested setting before claiming it will hold up in a busy hospital ward.
We also ran some informal trials with four older adults—ages 68, 72, 76, and 81—who interacted with the robot over two short sessions each. This was not a controlled study and we are not treating the observations as experimental data, but the pattern was consistent enough to note: all four engaged for longer with the sentiment-adaptive version than with a tone-neutral comparison condition, and two of them asked, unprompted, whether the robot had noticed something about how they were speaking. One said it "listened properly." That is exactly the effect [9] predicted, and it gives us confidence the sentiment module is doing real work.

VI. ETHICAL CONSIDERATIONS

Deploying a conversational device in a setting with elderly and potentially cognitively vulnerable users raises questions worth addressing directly.

Privacy was the first one the care facility raised with us, and it shaped a hard design constraint. Nothing a user says to BTA-Tron leaves the local network. There is no cloud logging of conversations, no third-party API receiving content beyond the STT call—and we are looking at local STT options to close that last gap too. Any version that adds health sensors would need explicit resident consent and a clearly communicated data retention policy before deployment.

The over-reliance risk is real. If an elderly person begins to treat interaction with a robot as a substitute for human contact rather than a supplement to it, that is not a good outcome—regardless of what the engagement scores show. We think deployment should always be accompanied by staff training that frames the robot as a gap-filler, not a replacement. The goal is to give residents more interaction overall, not to give human caregivers an excuse to visit less.
Finally, the emotion classifier is making inferences from words alone, without any deeper understanding of the person's situation. Someone who routinely understates distress—which is common in older adults—will not be correctly flagged. Human judgement needs to remain in the loop. The robot can notice patterns; it cannot replace clinical assessment.

VII. LIMITATIONS AND FUTURE WORK

A. What Did Not Work as Well as We Hoped

The sentiment classifier runs on transcribed text. When the speech-to-text step gets words wrong—which happens more often with quieter voices, or voices slowed by age or medication—the classifier is working with corrupted input. We noticed this in informal testing; a handful of perfectly clear emotional statements came back misclassified because of a transcription error upstream. Acoustic sentiment models that work directly on audio features would be more robust, and that is the most important technical improvement on our list.

The robot does not initiate conversation. It waits. In practice, the people who most need company are often the ones least likely to start talking to a machine. A proactive check-in feature—something that starts a conversation after a configurable silence, or checks in at a set time each day—would make the system substantially more useful and is high on the development plan.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505085|

The formal testing was done by us, in a lab. None of the performance numbers come from actual elderly users in a real care environment. The informal observations we described in Section V are encouraging but they are not a substitute for a properly controlled user study. That has to be the next serious step.

B. Planned Next Steps

Add a MAX30102 pulse oximeter, DS18B20 skin temperature sensor, and MPU-6050 inertial measurement unit so the robot can take basic health readings during conversation and log them to a caregiver dashboard.

Move the Flask server from a laptop to a Raspberry Pi 5 or Jetson Orin Nano, making the robot fully self-contained and eliminating the dependency on a nearby computer.

Build an acoustic sentiment classifier using MFCCs, pitch contour, and energy features as input to a small CNN—this would cut the STT error pathway entirely.

Add HC-SR04 ultrasonic sensors and a reactive path planner so the robot can navigate toward the user without needing manual driving.

Run a formal user study in a care facility with elderly participants, measuring engagement with the UTAUT2 instrument and loneliness outcomes with the UCLA-3 scale.

Integration of advanced natural language processing techniques for better and more adaptive interaction.
Addition of health monitoring sensors such as heart rate, temperature, and motion detection.
Development of autonomous navigation to allow the robot to move toward the user without manual control.
Implementation of cloud-based data storage for remote monitoring and data access.
Enhancement of the mobile application interface for improved accessibility and user experience.
Expansion of system capabilities to provide more intelligent and personalized assistance.

VIII. CONCLUSION

BTA-Tron does one thing that most assistive technology for elderly people does not: it talks back, and it pays attention to how the person is feeling when it does. Mean response time 4.1 s. Sentiment accuracy 87%. Motor command latency under one second. Total component cost under eighty dollars. No cloud required.

The design started from a conviction that loneliness is a health problem, not just a social one, and that a device designed to address it directly is doing something medically meaningful. The numbers and the informal observations both support that conviction. The next version—with physiological sensors, on-device processing, and autonomous navigation—is already in progress. But even this version makes the argument we wanted it to make: conversation-first assistive robotics works, it is affordable, and it fills a gap that no amount of wearable monitoring hardware will ever close.

IX. ACKNOWLEDGMENT

Thanks to the faculty of the Department of Electronics and Communication Engineering at Usha Mittal Institute of Technology for guidance, lab access, and patience with a project that involved a small robot driving around the workspace at inconvenient moments. Thanks also to the nursing staff at the care facility who gave us a candid and genuinely useful hour of their time at the start.

REFERENCES

[1]World Health Organization, “Ageing and health,” WHO Fact Sheet, Geneva, Switzerland, Oct. 2022. [link]
[2]T. G. Stavropoulos, A. Papastergiou, L. Mpaltadoros, S. Nikolopoulos, and I. Kompatsiaris, “IoT wearable sensors and devices in elderly care: a literature review,” Sensors, vol. 20, no. 10, p. 2826, May 2020. [link]
[3]J. Holt-Lunstad, T. B. Smith, M. Baker, T. Harris, and D. Stephenson, “Loneliness and social isolation as risk factors for mortality,” Perspect. Psychol. Sci., vol. 10, no. 2, pp. 227–237, Mar. 2015. [link]

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505085|

[4]A. Tapus, M. J. Mataríc, and B. Scassellati, "Socially assistive robotics," IEEE Robot. Autom. Mag., vol. 14, no. 1, pp. 35–42, Mar. 2007. [link]
[5]J. Broekens, M. Heerink, and H. Rosendal, "Assistive social robots in elderly care: a review," Gerontechnology, vol. 8, no. 2, pp. 94–103, 2009. [link]
[6]T.-C. Hsu and C.-J. Hsieh, "Socially assistive robots assisting older adults in an IoT and smart healthcare era," in Proc. 4th IEEE Eurasia Conf. IoT Commun. Eng. (ECICE), Oct. 2022, pp. 335–338. [link]
[7]N. Magnavita, I. Tripepi, and M. R. Di Prinzio, "Artificial intelligence, robots, and smart healthcare," Medicina, vol. 61, no. 1, p. 25, Jan. 2025. [link]
[8]L. Blavette, "Integrating a large language model into a socially assistive robot in a hospital geriatric unit," JMIR Hum. Factors, vol. 12, e58204, Mar. 2025. [link]
[9]H. Abdollahi, M. H. Mahoor, R. Zandie, J. Sewierski, and S. Qualls, "Artificial emotional intelligence in socially assistive robots for older adults: a pilot study," IEEE Trans. Affect. Comput., vol. 15, no. 2, pp. 692–705, 2024. [link]
[10]Z. Shi, E. Landrum, A. O'Connell, C. Groechel, R. A. Mataríc, and Q. Lim, "How can large language models enable better socially assistive human-robot interaction," arXiv:2404.00938, Apr. 2024. [link]
[11]M. Chita-Tegmark and M. Scheutz, "Assistive robots for the social management of health," Int. J. Soc. Robot., vol. 13, no. 7, pp. 1875–1892, 2021. [link]
[12]A. Rigaud, L. Moulin, and P. Jeanvoine, "Socially assistive robots in geriatric care: perceptions and barriers among healthcare professionals," Aging Clin. Exp. Res., vol. 36, p. 44, 2024. [link]
[13]S. Arunachalam, "Do socially assistive robots help elderly patients with frailty or mild cognitive impairment? A scoping review," Age Ageing, vol. 53, no. 4, afae071, Apr. 2024. [link]
[14]W. Khaksar, "Robotics in elderly healthcare: a review of recent work," arXiv:2301.04121, Jan. 2023. [link]
[15]J. Abdi, A. Al-Hindawi, T. Ng, and M. P. Vizcaychipi, "Scoping review on the use of socially assistive robot technology in elderly care," BMJ Open, vol. 8, no. 2, e018815, 2018. [link]
[16]V. Sheng, H. Zhang, and Q. Liu, "A WiFi-based indoor positioning and navigation system for assistive mobile robots," IEEE Access, vol. 10, pp. 21430–21441, 2022. [link]
[17]R. Lauria, G. Bugmann, T. Kyriacou, and E. Klein, "Mobile robot programming using natural language," Robot. Auton. Syst., vol. 44, nos. 3–4, pp. 171–180, 2003. [link]
[18]P. Wongpatikaseree, P. Kongthon, and A. Ratikan, "Developing a smart home system using IoT technology and home automation," in Proc. IEEE 4th Int. Conf. Comput. Commu

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE
SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details