International Journal of Advancing Research, Innovation and Entrepreneurship

Vol 1(2), DOI:10.82036/tr28me67, July 2025 *ISSN:2977-9332*

Autonomous Nanny Robot for Child Activity Recognition, Supervision and Caregiving

***[1]Kufre Esenowo Jack, [2]Mathew Emeka Nwanga, [3]Uzoma Gregory Okoro, [4]Ezekiel Gabriel Ufot,[5]Ezirim Kelechi Thankgod, [1]Emmanuel Chigozie Ani, [1]Samuel Akinwumi Adeniji, & [1]Samuel Adeniyi**

[1]Department of Mechatronics Engineering. Federal University of Technology Minna. Niger State. Nigeria.
[2]Department of Information Technology. Federal University of Technology Owerri. Imo State. Nigeria.
[3]Department of Mechanical Engineering. Federal University of Technology Minna. Niger State. Nigeria.
[4]Department of Electrical and Electronics Engineering, Akwa Ibom State University, Ikot Akpaden, Akwa Ibom State, Nigeria.
[5]Department of Mechatronics Engineering, Federal University of Technology Owerri, Imo State, Nigeria.
*Corresponding Author Email address: kufre@futminna.edu.ng

ABSTRACT

This study presents the design and implementation of an autonomous nanny robot aimed at enhancing childcare through intelligent monitoring and response. The robot employs machine learning, fuzzy logic, and real-time sensor technology to address key aspects of infant care, such as cry analysis, emotion recognition, and environmental safety. Cry modelling and analysis were performed in Mel-Frequency Cepstral Coefficients (MFCC). Emotion recognition was carried out based on a convolutional neural network (CNN), whereas the Environmental and physiological parameters, which included the body temperature, ambient temperature, and ammonia levels, were abstracted and evaluated by a fuzzy logic system to characterise the different levels of infant's comfort. The robot's design leverages a Raspberry Pi for processing and a Firebase for database management. The MFCC-based cry simulation demonstrated an accuracy of 77.87%, while the CNN achieved recognition accuracy of 85%. The fidelity of the models was tested based on precision-recall and F1 scores, and the results demonstrated fidelity in ensuring child safety, comfort, and caregiver convenience. This study further highlights the potential of integrating artificial intelligence and sensor-based technologies to redefine childcare solutions, offering a reliable and innovative approach to modern parenting challenges that allows caregivers to monitor conditions and receive alerts remotely.

Keywords: *Artificial Intelligence, Childcare Robot, Cry Analysis, Emotion Recognition, Environmental Monitoring, Fuzzy Logic.*

1 INTRODUCTION

Early childhood represents a critical period for physical and mental development, during which high-quality childcare can significantly influence various aspects of a child's growth trajectory (Chopra, 2015). Research underscores the role of quality early education and care in mitigating risks associated with challenging family environments, thereby fostering resilience and supporting positive developmental outcomes. High-quality childcare environments are shown to enhance social-emotional competencies, reducing conduct-related and emotional difficulties, particularly in children from high-risk backgrounds (Seiler et al., 2022). Optimistic and responsive interactions with caregivers contribute to the development of emotional regulation and prosocial behaviours, forming a foundation for healthy relationships in later life (Araujo et al., 2019). Early childhood also marks a period of rapid cognitive advancement, where supportive care environments play a role in the development of language, problem-solving abilities, and critical thinking (Joshi & Shukla, 2019). Studies reveal that children who experience high-quality care exhibit improved communication and cognitive abilities, establishing foundational skills that contribute to academic achievement (Burchinal et al., 2015). Quality childcare environments contribute to physical health by incorporating structured play and activities essential for motor skill development (Phillips & Lowenstein, 2011). Early engagement in varied physical activities supports overall health and wellness, laying a foundation for lifelong fitness habits (Donoghue et al., 2017).

For dual-income families, managing both work and childcare responsibilities presents substantial challenges, as parents must often navigate the demands of both domains. Research highlights the unique pressures placed on mothers, who frequently bear a more significant share of childcare duties, intensifying work-to-family conflict and, in some cases, necessitating compromises in their professional lives (Scheibling et al., 2023). The complex interplay between work and family responsibilities can lead to time constraints and heighten the risk of child neglect, particularly for working mothers striving to maintain a balance (Adem, 2023).

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

Effective time management is a significant challenge for parents, as they must allocate hours between work and childcare, which can lead to stress and burnout (Mortejo, 2024). High workplace demands often spill over into parenting roles, producing feelings of guilt and occasional absenteeism from both work and family responsibilities. (Shabunova & Leonidova, 2023). Establishing a reliable support network is essential for managing these dual responsibilities, as such networks provide much-needed assistance and emotional support (Adem, 2023). Families often employ adaptive strategies to manage these challenges. Flexible work arrangements, for instance, play a crucial role in accommodating family needs and reducing stress for dual-income parents (Luhr et al., 2022).

Traditional childcare methods present several limitations, including a lack of continuous monitoring, caregiver fatigue, and reliance on subjective judgment, all of which can adversely affect child development and well-being (Nivedha et al., 2020). Each of these issues reveals areas where conventional caregiving may fall short in promoting optimal outcomes. One prominent drawback is the absence of systematic oversight in traditional childcare, which can result in inconsistent and potentially harmful practices. Many caregivers rely on outdated methods that do not reflect current health guidelines, sometimes inadvertently exposing children to risks (Roy, 2020). Without professional guidance, traditional practices such as certain bathing or swaddling techniques may persist, even when they conflict with contemporary health recommendations (Morgan, 2021)

Caregiver fatigue is another significant issue within traditional childcare settings. Providers frequently experience high levels of stress and physical exhaustion, which can negatively impact the quality of care they deliver (Faulkner et al., 2016). Such fatigue often leads to diminished attentiveness and may exacerbate health issues among caregivers, further compromising the care provided to children. The perception of caregivers as merely "babysitters" can intensify their stress by reducing access to professional support and acknowledgement of their role's demands (Yinusa, 2024).

Subjective judgment, rooted in personal beliefs and experiences, also characterises much of traditional caregiving. This reliance on subjective knowledge introduces variability, as practices can differ widely among caregivers (Avetisyan & Konjoyan, 2023). In some cases, caregivers may prioritise cultural beliefs over evidence-based practices, creating inconsistency and potential health risks for children (Mizuki et al., 2020), (Pathak et al., 2022). While traditional childcare is often grounded in cultural heritage, its limitations underscore the need for enhanced education and support for caregivers. The subjective nature of traditional methods, along with the absence of continuous monitoring and the challenges of caregiver fatigue, points to the necessity for a more systematic and evidence-based approach to childcare (Lam et al., 2021).

Technological advancements have brought considerable changes to childcare, particularly through early interventions that improve child safety and allow for parental oversight. Innovations such as baby monitors, smart wearables, and IoT-based devices equip parents with tools to monitor their children's activities and enhance real-time safety measures (Pramod et al., 2024).

Smart wearable devices are a prime example of this trend. These devices, including wristbands and other smart gadgets, are equipped with sensors designed to notify parents when children approach potentially hazardous areas, such as kitchens or electrical appliances (Pramod et al., 2024). Through features such as GPS and Bluetooth tracking, these wearables provide real-time location updates, allowing parents to respond quickly in emergencies and continuously oversee their child's whereabouts (Shameema & L, 2023). IoT-based solutions further expand the capabilities of remote monitoring in childcare. For instance, IoT-enabled systems such as smart baby cradles enable caregivers to monitor vital signs and environmental conditions in real-time, enhancing both safety and convenience (Journa, 2023). Beyond immediate monitoring, these systems often provide data-driven insights that allow caregivers to refine childcare practices and respond proactively to potential issues (Dumpalwar, 2024).

Despite advances, current childcare technologies exhibit notable limitations, which can restrict their effectiveness in providing optimal caregiving environments. A primary concern is the lack of comprehensive supervision capabilities. Many of these technologies do not fully support the breadth of oversight needed for high-quality childcare, resulting in environments that may fall short of fostering well-rounded development (Chappara et al., 2023). Another significant limitation is the absence of autonomous decision-making that caters to individual child needs. Most existing systems are unable to make nuanced, informed decisions critical for personalised care, limiting their ability to respond dynamically to the varying requirements of each child (Singh et al., 2023). Furthermore, the inability to replicate responsive caregiving interactions remains a key shortcoming. These technologies are yet to capture the subtle, adaptive responses of human caregivers, which are essential for supporting emotional and cognitive growth (Sumathi et al., 2023).

The primary aim of this research is to develop an autonomous nanny robot designed to integrate real-time child activity recognition, supervision, and caregiving functions. This innovative robot seeks to assist parents and caregivers by autonomously monitoring and responding to a child's activities, offering a more interactive and personalised form of care. By integrating advanced technologies, the robot addresses the limitations of traditional childcare practices and existing childcare technologies, ultimately providing continuous, tailored supervision.

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

To accomplish this purpose, this research includes several key objectives. The first is to fabricate the nanny robot along with a stable carriage system, ensuring a secure, mobile platform capable of navigating domestic spaces safely. Another objective is to design an intelligent control system for managing the robot's movement, obstacle detection, and interactions with the child. This control system incorporates various sensory inputs to enable autonomous navigation and response. Additionally, the second objective is to implement cry and sound analysis capabilities, enabling the robot to distinguish between sounds and recognise a child's cries, particularly in situations that may require immediate attention or intervention. Furthermore, the third objective is the integration of emotion recognition technology, which will allow the robot to interpret the child's emotional state and respond appropriately based on observed cues. The last objective is the development of a remote monitoring system, providing parents and caregivers with a user-friendly interface to observe and interact with the nanny robot's activities in real-time, along with notifications and alerts when necessary.

The importance of this research is significant in the field of automated caregiving and childcare technology. By addressing critical gaps in traditional childcare, such as continuous monitoring and the need for real-time responses to a child's emotional and physical states, the nanny robot provides a solution tailored to modern family dynamics, particularly in dual-income households and situations requiring extended work hours. This autonomous robot not only supports parents with enhanced safety and engagement for their children but also contributes to advancing child-centred AI research, emotional analysis, and real-time monitoring in childcare settings. Ultimately, the research aims to improve early childhood development and safety through a comprehensive caregiving system, potentially shaping the future of robotics in childcare and providing a vital support system for parents and caregivers.

The development of this autonomous nanny robot directly aligns with the needs of modern, busy parents who seek safe, reliable, and responsive childcare solutions. In today's fast-paced world, parents often juggle demanding work schedules and family responsibilities, making it challenging to provide continuous supervision and support for their children. This robot addresses these concerns by offering real-time monitoring and interaction, ensuring that children are safe and engaged while parents are occupied with other commitments. Its ability to analyse child activity, respond to cries, and recognise emotions enables it to provide a level of care that adapts to individual needs, which is essential for the emotional and cognitive development of young children.

This autonomous nanny robot has the potential to significantly improve the consistency of care provided to children. Reducing reliance on human nannies can help address common issues such as caregiver fatigue, subjectivity in caregiving practices, and the variability of care quality. The robot's systematic approach to monitoring and responding to children creates a more stable environment that supports healthy development. This technology not only promotes better outcomes for children but also empowers parents by providing a reliable alternative that enhances their peace of mind and facilitates a balanced approach to family life.

2 SUMMARY OF RELEVANT LITERATURE

The concept of autonomous robots in caregiving is rapidly gaining traction, particularly as a means to assist children and the elderly (Ashwini & Rathnakara, 2024). These robots are designed to leverage advanced technologies to enhance the quality of care provided, alleviate feelings of loneliness, and offer personalized support tailored to the unique needs of their users. With the growing demand for innovative caregiving solutions, the development of autonomous robots represents a significant advancement in the intersection of robotics, artificial intelligence, and human-centred design (Hsu & Šabanović, 2024).

The Autonomous Nanny Robot, developed for caregiving applications, embodies this technological evolution by integrating various advanced features. These include sensor technology for environmental monitoring, artificial intelligence (AI) for interaction and decision-making, and mobility capabilities for seamless navigation in dynamic environments (Duthie et al., 2024). By combining these elements, the robot aims to provide an intelligent, adaptive, and user-friendly solution to caregiving challenges. This review seeks to contextualise the research within the existing body of research, highlight its contributions, and identify gaps that it seeks to address (Gongor & Tutsoy, 2024).

Analysing baby cries relies on advanced signal processing and machine learning techniques. MFCC are a key feature for cry classification, providing compact audio representations (Qiu et al., 2024; Krishna et al., 2024). Spectrograms also enhance analysis by revealing temporal patterns (Jindal et al., 2021). Hybrid feature sets, combining MFCC with others like Mel Spectrogram and Tonnetz, have improved classification accuracy, exceeding 94% in some studies (Qiu et al., 2024).

Machine learning models such as K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs) effectively analyse cry patterns (Ji et al., 2021; Jindal et al., 2021). Dimensionality reduction techniques like Principal Component Analysis (PCA) further optimise feature space while maintaining accuracy (Krishna et al., 2024).

Recognising emotions in infants through facial expressions and cries is critical for effective caregiving and early childhood development (Das et al., 2024). Advanced neural network models like CNNs, ResNet, and attention-based architectures, such as SE-ResNet-Transformer and ResNet-BiLSTM, have improved emotion classification accuracy by

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

leveraging their robust feature extraction and classification capabilities (Mishra et al., 2024). These systems help caregivers respond to infants' emotional needs and support mental health assessments and human-computer interactions (Shen et al., 2024)

However, challenges persist, including the lack of publicly available datasets for infant speech emotion recognition, which limits model development (Shen et al., 2024). Additionally, balancing model simplicity and effectiveness is crucial, as overly complex architectures demand significant computational resources and training time (Gupta et al., 2024). Variability in infant expressions and cries further complicates achieving high accuracy. Addressing these limitations is essential to developing reliable systems that enhance caregiving and foster a better understanding of infant emotions (Li et al., 2024).

Sensors play a vital role in childcare robots by ensuring environmental safety and effective physiological monitoring. Temperature and humidity sensors, such as the DHT22, monitor and regulate ambient conditions, enhancing the comfort and safety for infants (D'Sa et al., 2024). Gas sensors like the MQ-135 detect harmful gases, maintaining air quality and alerting caregivers to potential hazards (S Raju & T, 2022). Additionally, physiological monitoring, such as tracking body temperature, ensures comprehensive health assessments, enabling timely responses to deviations from normal ranges (Alva et al., 2024).

Integrating environmental and physiological data enhances decision-making in robotic systems. Decision-theoretic approaches allow robots to autonomously optimise data collection, maintaining safe and adaptive childcare environments (Sung et al., 2023). However, challenges in accuracy and reliability remain, necessitating advancements in sensor technology and algorithms to improve robotic efficacy and reliability in childcare settings (Bogue, 2023).

Mobile applications are essential in childcare robots, providing real-time monitoring, interaction, and control capabilities. These apps enable caregivers to manage robot movements, access live video feeds, and receive timely alerts about potential risks, thereby enhancing child safety (Alva et al., 2024). They also support educational activities by offering interactive games and learning tools tailored to a child's developmental stage. Examples include the COPPER Babysitter Robot, which integrates manual and automatic control, video streaming, and safety alerts, and an IoT-based child safety robot with live streaming and hazard notifications like gas leak alerts (Alva et al., 2024).

Mobile apps rely on effective communication protocols such as LINE notifications for immediate risk alerts and webhooks for cost-efficient, network-based connectivity. Real-time video feeds and interactive interfaces provide intuitive data visualisation, enabling caregivers to monitor and engage seamlessly with children. However, while these integrations offer significant benefits, addressing privacy and security concerns is crucial to protecting sensitive data and ensuring secure communication channels.

The Autonomous Nanny Robot exemplifies the potential of robotics to transform caregiving by integrating advanced sensor systems, AI capabilities, and mobility solutions. This literature review aims to delve deeper into these components, examining existing research to identify how the research builds upon established knowledge and addresses gaps to create a more adaptive, safe, and effective caregiving solution.

3 METHODOLOGY

The methodology for developing this autonomous nanny robot integrates all aspects of design, fabrication, control systems, and implementation to address child activity recognition, supervision, and caregiving needs. This comprehensive approach ensures that the system is reliable, responsive, and effective in meeting modern childcare requirements. The methodology emphasises the fabrication of the robot and its carriage system, the design of a robust control system, and the implementation of sensory, analytical, and monitoring features.

3.1 STRUCTURAL DESIGN AND IMPLEMENTATION OF THE NANNY ROBOT

This research commenced with the design and fabrication of the nanny robot and its carriage system. Conceptual modelling was carried out using Computer-Aided Design (CAD) software to ensure an ergonomic and structurally sound design capable of accommodating sensors and electronic components. The physical frame was constructed from lightweight aluminium to provide durability without excess weight, while the body was 3D-printed using PLA or ABS filaments. Post-processing techniques, including sanding and finishing, were applied to achieve a smooth surface, ensuring safety and aesthetic appeal. Finite Element Analysis (FEA) was performed to evaluate structural integrity under various load conditions, analysing stress distribution, displacement, and strain to validate the design's durability under both static and dynamic loads, as illustrated in Figure 1.

Following the structural design, the electronic circuitry of the system was developed, with the ESP32 microcontroller serving as the central processing unit due to its low power consumption and robust wireless capabilities. The schematic design, shown in Figure 2, detailed the connections between the ESP32 and the various sensors, including the DHT11 temperature sensor, non-contact temperature sensor, MQ-135 gas sensor for ammonia detection, and other components such as resistors, capacitors, and voltage regulators for stable operation. The sensors were strategically positioned to optimise performance; the temperature sensor was placed for accurate ambient temperature readings, while the gas sensor

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

was positioned to detect harmful gases in the infant's environment.

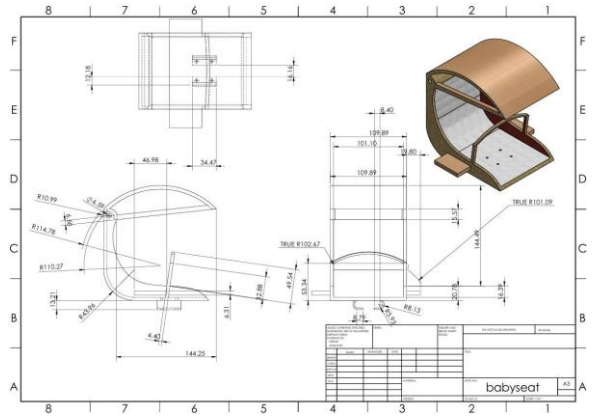


Figure 1: Render of the baby carriage system

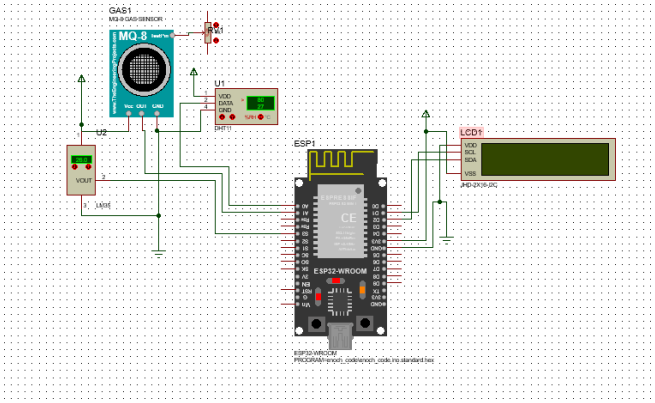


Figure 2: Circuit Schematic of the Infant Monitoring System

3.2 CONTROL SYSTEM FOR THE NANNY ROBOT

The system's control logic was developed using the Robot Operating System (ROS 2) middleware to unify the robot's operations and manage decision-making across subsystems. The Raspberry Pi 4 served as the core microcontroller, executing all logic related to navigation, locomotion, and sensory data processing. The system employed a monocular camera, ultrasonic and infrared distance sensors, a microphone, and an Inertial Measurement Unit (IMU) to gather environmental data, which was processed for decision-making. ROS 2 nodes were configured to handle tasks such as navigation, joint motion, audio analysis, and video analysis for emotion recognition. Figure 3 illustrates the block diagram of the control subsystems.

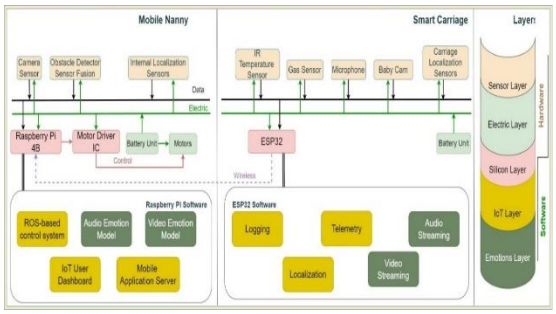


Figure 3: Block diagram of the Childcare Robot Control Subsystems

Simulations were conducted using the Webots software to test and refine the control system before hardware implementation. The PR2 robot model, obtained from the Webots repository and modified to match the nanny robot's sensor and actuator configuration, was used to simulate various scenarios within an apartment environment. These simulations included obstacle avoidance, navigation, and task execution, such as rocking the child's carriage. Figure 4 presents a side-by-side comparison of the PR2 model and the nanny robot. The simulation phase identified and resolved issues using ROS 2 introspection tools and the PlotJuggler package.

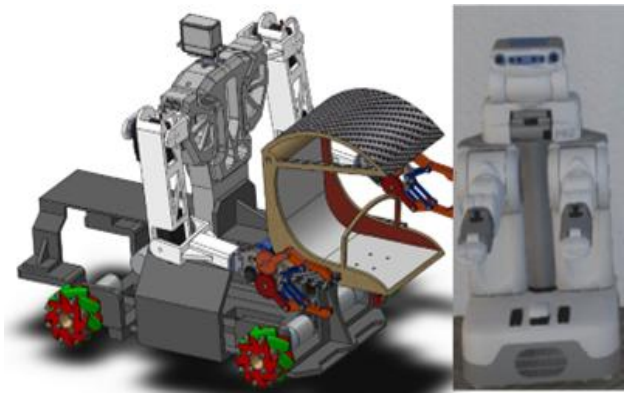


Figure 4: Side-by-side Comparison of Nanny Robot (left) and PR2 (right)

For navigation, the system implemented Simultaneous Localization and Mapping (SLAM) using the SLAM Toolbox package in ROS 2. The monocular camera input was utilised to create environmental maps and locate the child's carriage. Once the carriage was identified, the robot navigated to its position using the Nav2 framework, which employed behaviour trees for path planning and obstacle avoidance. The behaviour tree, depicted in Figure 5, outlined decision-making processes, including recovery behaviours for navigation conflicts. Upon reaching the carriage, the robot used preprogrammed joint motions to dock and rock the carriage.

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

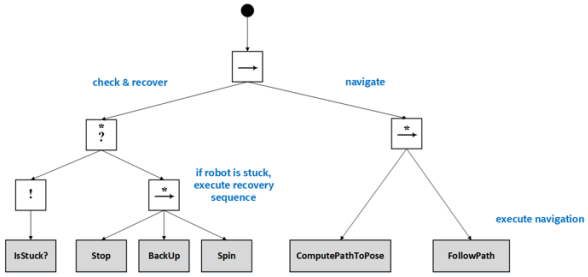


Figure 5: Behaviour Tree Used for Robot Navigation

The control system also featured an assistive loop to detect and respond to a baby's cries. The robot remained on standby, analysing audio in 30-second intervals using an audio analysis model. Figure 6 illustrates the flow chart for these control operations, which were managed using a lifecycle node system to coordinate subsystems and sensor fusion for odometry using the Extended Kalman Filter implementation within Nav2.

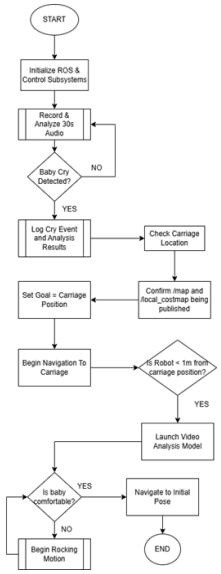


Figure 6: Flow Chart of Control Operations for the Nanny Robot

Locomotion was achieved through a four-wheeled differential drive system controlled by a ROS 2 Differential Drive Controller node. The controller converted navigational instructions into motor speed commands and calculated odometry information for localisation. Joint control for the robot's arms was implemented via a manually programmed node using the PiGPIO Pin Factory to minimise servo jitter during movements. Joint calibration ensured accurate motion execution, and configurations were defined in a YAML file for precise actuation.

3.3 CRY AND SOUND ANALYSIS FOR THE NANNY ROBOT

This research uses the MFCC technique with supervised learning and a CatBoost model to classify baby cries into three classes: hunger, discomfort, and tiredness. The training dataset was composed of 457 cries from the "donate-a-cry-corpus" and 155 locally sourced cries, with a focus on these three particular classes due to dataset imbalance. MFCC is used for feature extraction to convert the infant cries into a processable form suitable for Raspberry Pi. The extracted features are then fed into the CatBoost model for training. The model was tested with randomly selected samples to evaluate its accuracy and effectiveness before being deployed on the Raspberry Pi 4 to analyse infant cries in real-time for interpreting their needs.

The data source for this research is the "Donate-a-cry Corpus," a publicly available dataset containing five types of infant cries—belly pain, hunger, burping, discomfort, and tiredness—based on the Dunstan Baby Language theory. The original dataset contained 488 samples with varying class distributions: 24 for belly pain, 58 for discomfort, 382 for hunger, 24 for tiredness, and 16 for burping. To address class imbalances, additional audio samples were collected from hospital recordings and YouTube, raising the sample counts for discomfort to 234, hunger to 210, and tiredness to 168. Due to insufficient sample sizes for belly pain and burping, these two categories were excluded from this research.

The supervised learning method employed the MFCC technique and the CatBoost model for feature extraction and model training, with the dataset focusing on discomfort, hunger, and tiredness. Table I shows the data sample distribution across the three classes of interest. The feature extraction using MFCC converts the infant cry into distinct features that the Raspberry Pi then processes. The trained CatBoost model was tested using randomly selected samples, and the results were evaluated for effectiveness before deployment onto the Raspberry Pi 4.

TABLE I: CLASSES AND SAMPLE DATA USED

Donate a cry-corpus	Discomfort	Hungry	Tired
Original	27	382	24
Locally sourced	207	0	144
Reduction of cry sample	0	172	0
Total cry sample	207	210	144

The model's specifications involve the CatBoost algorithm, which constructs models by sequentially adding weak learners to minimise a loss function. CatBoost handles categorical features through target-based encoding and uses ordered boosting to prevent overfitting. This methodology made CatBoost an ideal choice for this research. After training, the model was deployed onto the Raspberry Pi 4, which is connected to the microphone to analyse infant cries

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

and predict the underlying needs based on their emotional state.

The system design for analysing infant cries using AI includes deploying the trained model on the Raspberry Pi 4 and connecting the microphone to the device. Figure 7 presents the circuit diagram of the sound analysis system, which illustrates the system's components and their interconnections.

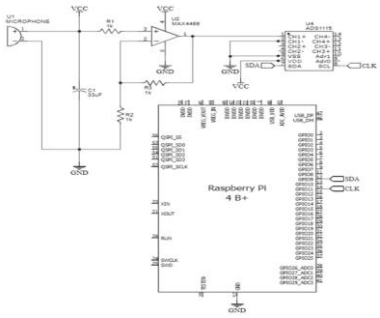


Figure 7: Circuit Diagram for the sound analysis system for the Nanny Robot

3.4 Emotion Recognition System for the Nanny Robot

Additionally, this research incorporates an infant emotion recognition system which utilises artificial intelligence by combining deep learning, sensor networks, and fuzzy logic to create an effective caregiving robot. The Raspberry Pi 4 acts as the central processing unit, handling data acquisition and processing. It is equipped with a camera module for facial emotion detection, a temperature sensor for ambient monitoring, a gas sensor to detect ammonia levels from baby diapers, and an infrared temperature sensor.

This emotion recognition system uses a dataset of 1,451 images, consisting of 134 images from the Troms Database and 1,317 images sourced from YouTube videos. These images are used to train the system to recognise an infant's three primary emotional states: happy, neutral, and sad. The system is designed for real-time monitoring through a web interface, alerting caregivers to the child's emotional state.

The system flow chart (Figure 8) outlines the process of the infant emotion recognition system. When activated, the system captures the infant's image and checks for a face. If a face is detected, the system uses the trained model to assess the infant's emotional state based on facial expressions. If no cry is detected, the system continues monitoring by capturing additional images. If the infant is crying, the system collects environmental data for context and uses fuzzy logic to assess the cry's intensity and environmental factors. Finally, the system notifies the caregiver with information about the infant's emotional state.

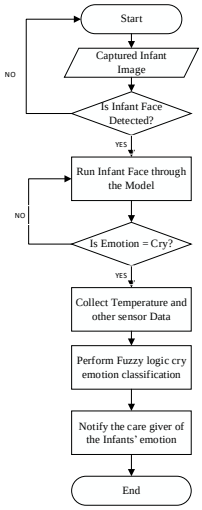


Figure 8: Flow chart for emotion recognition

The data pre-processing pipeline involves resizing images to 299×299 pixels and splitting the dataset into training and testing sets with an 80:20 ratio. The training process, as illustrated in Figure 9, involves using the Inception V3 pre-trained convolutional neural network (CNN) model. The model is enhanced with a Squeeze and Excitation (SE) attention mechanism and a Convolutional Block Attention Module (CBAM) to focus on key facial features, such as the eyes and mouth. The output layer uses a SoftMax activation function for classifying the infant's emotional state.

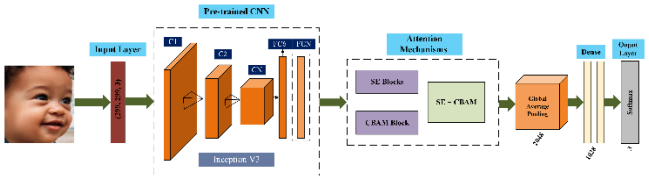


Figure 9: Model Architecture for Infant Emotion Recognition

3.5 Remote Monitoring System for the Nanny Robot

A remote monitoring system is integrated with the autonomous nanny robot to provide real-time data collection and remote monitoring for the child's health and safety. The system's design ensures that caregivers can continuously track vital environmental factors such as temperature, humidity, air quality, and the child's body temperature from a distance, enhancing the robot's caregiving capabilities.

The hardware setup incorporates various sensors, each playing a crucial role in ensuring the well-being of the child. The Raspberry Pi 4B is used as the central processing unit, handling data collection and communication with the cloud via Wi-Fi or Ethernet. It processes inputs from the sensors and sends them to the Firebase Real-time Database for remote monitoring. The TS118-3 non-contact infrared temperature sensor monitors the child's body temperature,

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

while the MQ-135 gas sensor detects ammonia levels in the air, alerting caregivers when thresholds are exceeded. The DHT22 sensor measures the ambient temperature and humidity, ensuring a comfortable environment for the child. These sensors are carefully selected to ensure continuous monitoring and to alert caregivers of any irregularities that may require attention.

Python scripts are written to control the sensors, process the collected data, and send it to Firebase via HTTPS, ensuring secure communication between the Raspberry Pi and the cloud. The Firebase Real-time Database stores the sensor data, providing real-time synchronisation and allowing caregivers to access the information remotely through the mobile app. The app, developed using MIT App Inventor, presents the sensor data in a user-friendly interface, making it easy for caregivers to monitor the child's environment and health status. The process is explained in figure 10.

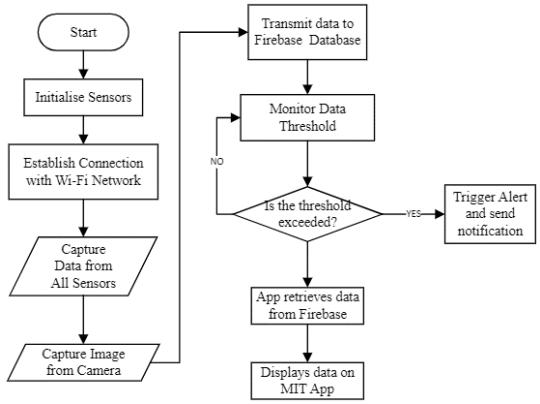


Figure 10: Flowchart of the remote monitoring system

The communication between the Raspberry Pi and Firebase uses HTTP/HTTPS protocols to ensure data security. The system also implements security measures to safeguard the stored data and ensure only authorised users can access the information. In case of abnormal readings, such as temperature or humidity levels crossing predefined thresholds, the system triggers alerts, notifying caregivers through the mobile app. These notifications provide timely information, enabling caregivers to take immediate action if necessary.

4 RESULTS AND DISCUSSION

The results of this research are presented to align with the objectives, focusing on the design, implementation, and control of the baby carriage system.

4.1 STRUCTURAL DESIGN AND IMPLEMENTATION OF THE NANNY ROBOT

The Finite Element Analysis (FEA) of the baby carriage design provided critical insights into its mechanical performance under various load conditions, focusing on stress distribution, displacement, and strain. The von Mises stress distribution, depicted in Figure 11, highlights a colour-coded scale ranging from 1.958e+02 N/m² (dark blue) to 7.377e+06 N/m² (dark red), showing varying stress levels within the structure. The analysis confirmed that the stress remained within safety limits, as it did not exceed the material's yield strength. Critical stress zones, particularly around load application points, were identified but deemed structurally sound under the given conditions.

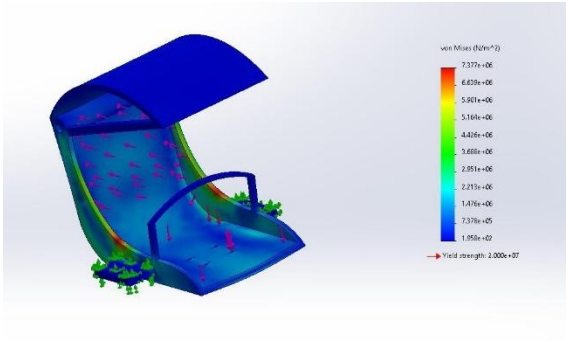


Figure 1: Stress Distribution of the Structural Component under Applied Load Displacement result

Figure 12 shows displacement analysis revealing the structural response to applied loads. Maximum displacement occurs in the upper portion of the structure (approximately 2.177 mm). The displacement reduces significantly near constraints, maintaining the design's stability.

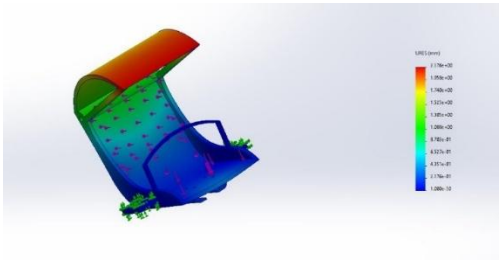


Figure 2: Displacement distribution of the structural component under applied load

Strain distribution analysis, illustrated in Figure 13, indicated varying strain levels across the structure, with maximum strain concentrated in areas experiencing greater deformation under applied loads. These findings confirmed the structural integrity of the design under the given conditions.

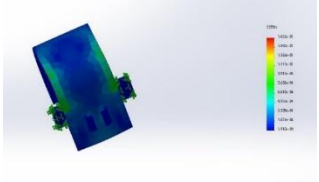


Figure 3: Strain distribution of the structural component under applied load

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

The implementation phase of the baby carriage system involved assembling and integrating components to create a functional prototype. This process included constructing the frame, installing electronic components, and integrating sensors for a comprehensive monitoring system. The system met safety standards and operated effectively under real-world conditions, showcasing its readiness for evaluation and testing in diverse scenarios.

4.2 Result and Discussion for Control System for the Nanny Robot.

The relevant software dependencies were installed for the control system, and the control system was created as a package in ROS2. Configuration files incorporated variables used by the ROS2 packages and the robot simulation. A launch file enabled the system to be brought up with a single command, with options for independent launches of specific components. The ROS2 introspection tools in RQt verified internode communication and confirmed successful robot initialisation. The monocular camera model's output was inspected within the simulation environment, as shown in Figure 14.



Figure 4: Initial Camera Output Showing Simulated Apartment Building

The navigation system allowed the robot to form a map of the environment, as depicted in Figure 15, by sending velocity commands via the keyboard. Obstacle detection functionality was verified through local and global cost maps, ensuring accurate navigation. The robot's decision-making was tested by directing it to navigate to random points on the map, demonstrating adherence to its behaviour tree.

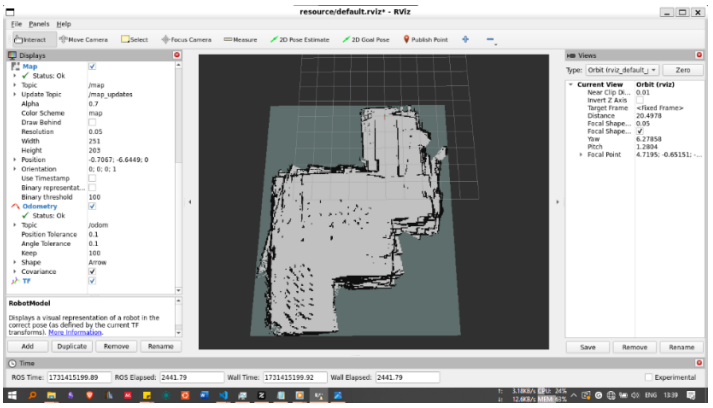


Figure 5: Map of Navigable Area Within the Simulated Apartment Building

Velocity commands during navigation cycles were monitored, as shown in Figures 16 and 17, confirming expected behavior with commands aligned to specific axes.

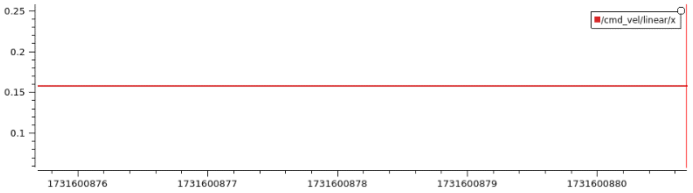


Figure 6: Graph of Linear Velocity Commands Given to the Locomotion System by the Real Time Navigation System in m/s

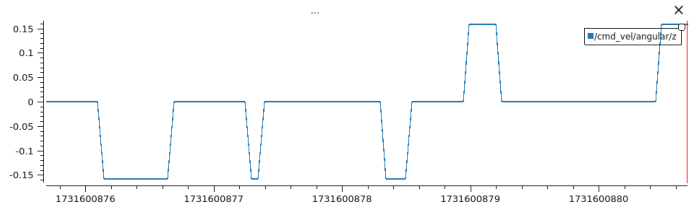


Figure 7: Graph of Angular Velocity Commands Given to the Locomotion System by the Real Time Navigation System in m/s

Real-time changes in robot heading and position were analyzed using IMU sensor data. The robot's orientation and angular velocity over time were measured, with angular velocity responses lagging by approximately 0.4 seconds, as shown in Figures 18 and 19. Linear acceleration during navigation, depicted in Figure 20, confirmed stable performance under commands.

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

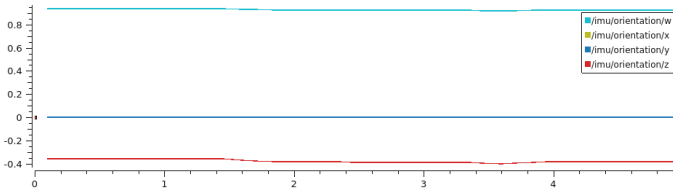


Figure 8: Graph of Robot Orientation on the Map as Measured by the IMU Sensors

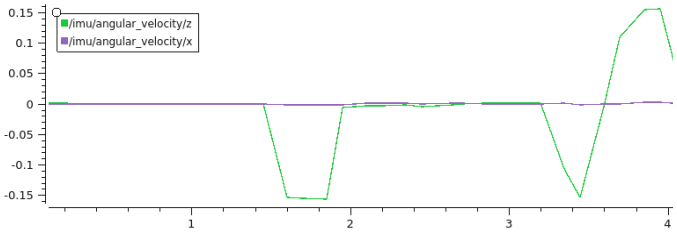


Figure 9: Graph of Robot Angular Velocity During Locomotion in m/s

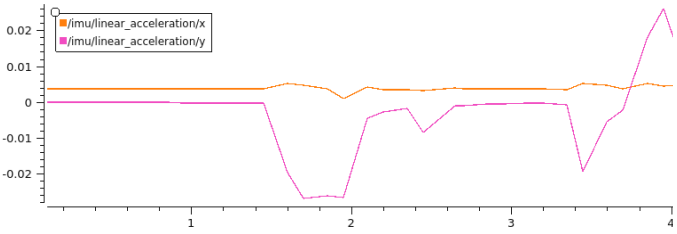


Figure 10: Graph of Robot Linear Acceleration During Locomotion in m/s^2

The control system package was built and exported with a launch file for automatic startup, and the system was configured for boot on a Raspberry Pi. Navigation was hosted on an external Ubuntu PC within the same local network. Joint controllers were tested on the hardware, and all joints were successfully initialized and actuated, confirming the system's operational integrity.

4.3 Result and Discussion for Cry and Sound Analysis for the Nanny Robot

For the baby cry analysis, the features from the acquired dataset were extracted using the MFCC technique and subsequently fed into the model. To ensure accurate learning, the model was trained with a learning rate of 0.1, an iteration value of 1000, a depth of 15, and a random state of 42. The training process yielded notable results, as depicted in Figure 21.

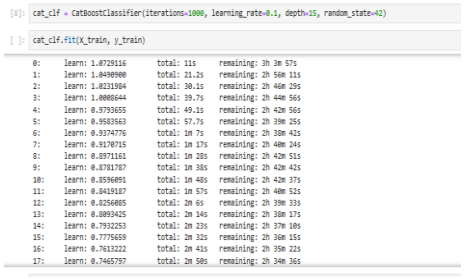


Figure 11: Training process obtained from the model

In evaluating the model's performance, key metrics such as precision, recall, and F1-score were employed alongside a confusion matrix to represent the testing outcomes. The evaluation results are summarised in Table 2, showing an accuracy of 77.87%.

TABLE 1: METRIC RESULTS FROM MODEL TESTING

	Precision	Recall	F1-score	Support
1	0.85	0.74	0.79	46
2	0.71	0.93	0.80	42
3	0.81	0.65	0.72	34
Accuracy			0.78	122
Macro avg	0.79	0.77	0.77	122
Weighted avg	0.79	0.78	0.78	122

The confusion matrix, shown in Figure 22, further illustrates the performance, providing insight into the distribution of predictions.

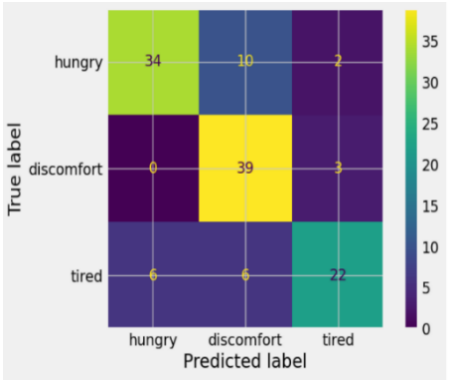


Figure 12: Confusion matrix of the model testing

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

4.4 Result and Discussion of Emotion Recognition System for the Nanny Robot

For the facial emotion recognition task, the dataset included 1,064 images of infants displaying three emotional states: happy, neutral, and sad. The data distribution, as shown in Figure 23, consisted of 518 happy images (48.87%), 51 neutral images (4.79%), and 495 sad images (46.52%). The training process spanned 30 epochs with batches of 64 images, achieving a peak training accuracy of 0.86 and validation accuracy of 0.54 at epoch 18. Training and validation loss reached their lowest points at 0.85 and 0.43, respectively, by epoch 12.



Figure 13: Sample Infant Faces with Three Emotional States

Testing the facial emotion recognition model on 267 images (130 happy, 13 neutral, and 124 sad) resulted in an overall accuracy of 85%. The confusion matrix in Figure 24 highlights the model's ability to correctly predict emotions, with a precision of 0.94, recall of 0.78, and F1-score of 0.85 for happy emotions; a precision of 0.86, recall of 0.46, and F1-score of 0.60 for neutral emotions; and a precision of 0.80, recall of 0.98, and F1-score of 0.88 for sad emotions. The ROC-AUC curve, depicted in Figure 25, provides additional insight into the model's performance, demonstrating its robustness in distinguishing between emotional states.

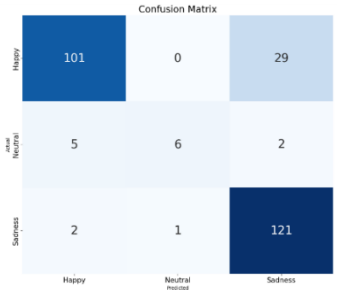


Figure 14: Infant Emotion Confusion Matrix for the Test set

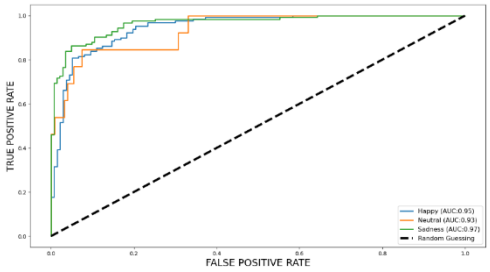


Figure 15: ROC-AUC curve of the infant emotion recognition model

Fuzzy logic was applied for cry emotion classification using membership functions for body temperature, external temperature, and ammonia concentration. The body temperature membership function (Figure 26) scategorised temperatures into "Low," "Normal," and "High," while the external temperature membership function identified "Cold," "Comfortable," and "Hot" conditions.

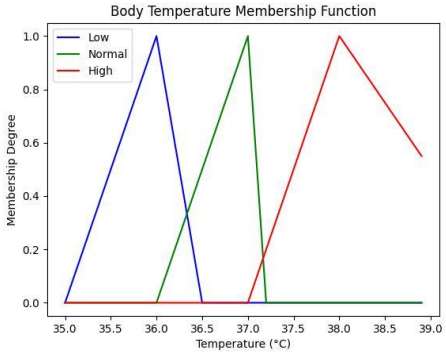


Figure 16: Infant Body Temperature Membership Function

4.5 Result and Discussion for Remote Monitoring System for the Nanny Robot

The remote monitoring app interface as shown in Figure 27 was designed for caregivers to access critical data about the child's environment and well-being in real time. The landing page provided a simple and intuitive entry point, ensuring easy navigation. The sensors data displayed comprehensive real-time data, including temperature, humidity, body temperature, and air quality, in a user-friendly format. The live feed page offered a continuous video stream from the robot's camera with minimal latency, enabling caregivers to visually monitor the child and their surroundings. This holistic approach to remote monitoring represents a significant advancement in childcare technology, improving safety and providing caregivers with confidence and peace of mind.

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

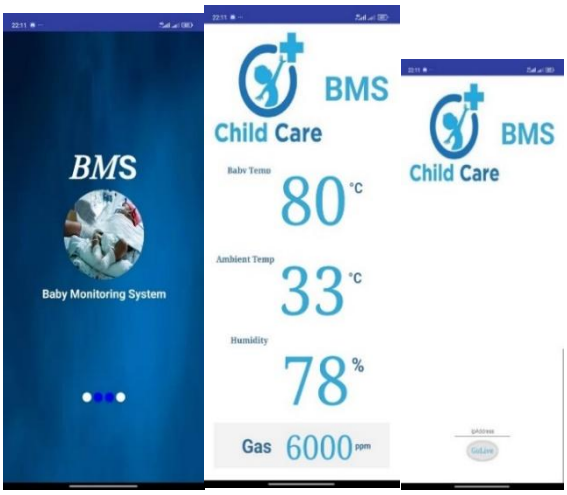


Figure 17: From L-R: Landing Page, Sensor Monitoring Page, Live Stream Page

5 CONCLUSION

This research presents the successful development and implementation of an autonomous nanny robot equipped with advanced monitoring and predictive capabilities to enhance childcare. By integrating cry analysis and facial emotion recognition using machine learning techniques, the system demonstrated the ability to interpret an infant's needs and emotional states with high accuracy. The cry analysis model achieved a precision of 0.79, recall of 0.77, and an overall accuracy of 77.87%. In comparison, the facial emotion recognition model recorded an impressive test accuracy of 85%, effectively classifying happy, neutral, and sad emotions. The incorporation of fuzzy logic further enhanced the robot's decision-making ability by assessing environmental and physiological parameters, such as body temperature, ambient temperature, and ammonia concentration, ensuring the comfort and safety of the child. The seamless integration of sensors for real-time environmental and health monitoring, including DHT22, TS118-3, and MQ-135, validated the robot's reliability in detecting and responding to dynamic changes in the child's surroundings. The remote monitoring application, featuring intuitive interfaces for live video streaming and real-time sensor data, provided caregivers with an efficient and user-friendly platform for oversight. This innovative approach combines advanced artificial intelligence, fuzzy logic, and robust sensor integration to redefine childcare solutions. Future work will focus on expanding the robot's capabilities, including the integration of additional emotional states, enhanced mobility features, and multilingual support, to address a broader range of scenarios. The findings underscore the potential of autonomous systems in revolutionizing caregiving, offering safer, smarter, and more effective childcare solutions.

REFERENCES

Adem, B. S. (2023). The Intersection of Parenting and Employment: A Qualitative Exploration of the Experiences of Working Parents. *Research Square*. https://doi.org/10.21203/rs.3.rs-3772462/v1

Alva, R. A., Herrera Poma, A. L., Perez Villa, V. E., & Peña, F. W. Z. (2024). The COPPER Babysitter Robot, a Childcare Monitoring System from the First year of Age. In The COPPER Babysitter Robot, a Childcare Monitoring System from the First year of Age (2024 ed., Vol. 1130, pp. 83-95). *Springer Nature*. https://doi.org/10.1007/978-3-031-50381-8_11

Araujo, M. C., Dormal, M., & Schady, N. (2019). Childcare Quality and Child Development. *Journal of Human Resources*, 54(3), 656–682. https://doi.org/10.3368/jhr.54.3.0217.8572R1

Ashwini, G., & Rathnakara, S. (2024). A Novel Approach to Elderly Care Robotics Enhanced With Leveraging Gesture Recognition and Voice Assistance. *2024 Asia Pacific Conference on Innovation in Technology (APCIT)*, 1–6. https://doi.org/10.1109/APCIT62007.2024.10673657

Avetisyan, M., & Konjoyan, H. (2023). A Comparative Analysis between Traditional and Alternative Early Childhood Education, Such as the Maria Montessori Method, Reggio Emilia Pedagogical Approach. *Armenian Journal of Special Education*, 7(2), 65–77. https://doi.org/10.24234/se.v7i2.8

Bogue, R. (2023). The role of robots in environmental monitoring. *Industrial Robot: The International Journal of Robotics Research and Application*, 50(3), 369–375. https://doi.org/10.1108/IR-12-2022-0316

Burchinal, M., Magnuson, K., Powell, D., & Hong, S. S. (2015). Early Childcare and Education. *In Handbook of Child Psychology and Developmental Science* (pp. 1–45). Wiley. https://doi.org/10.1002/9781118963418.childpsy406

Chappara, M., Ribeiro, F., & Metrôlho, J. (2023). UNDERAGE CITIZENS MONITORING APPLICATIONS - A REVIEW OF THE STATE OF THE ART AND GUIDELINES FOR FUTURE IMPLEMENTATIONS. *International Journal of Engineering Science Technologies*, 7(4), 1–18. https://doi.org/10.29121/ijoest.v7.i4.2023.524

Chopra, G. (2015). Early Childhood Care and Education: Right to Survival and Development. In Child Rights in India (pp. 45–76). *Springer India*. https://doi.org/10.1007/978-81-322-2446-4_3

Das, C., Nanda, D., Chatterjee, S., & Dutta, D. (2024). Comparative Analysis of Neonatal Facial Expression from Images using Deep Learning. 2024 *15th*

Vol 1(2), DOI:10.82036/tr28me67, July 2025 *ISSN:2977-9332*

International Conference on Computing Communication and Networking Technologies (ICCCNT), 1–7. https://doi.org/10.1109/ICCCNT61001.2024.10725190

Donoghue, E. A., Lieser, D., DelConte, B., Donoghue, E., Earls, M., Glassy, D., Mendelsohn, A., McFadden, T., Scholer, S., Takagishi, J., Vanderbilt, D., & Williams, P. G. (2017). Quality Early Education and Child Care From Birth to Kindergarten. *Pediatrics*, 140(2). https://doi.org/10.1542/peds.2017-1488

D'Sa, R. J., Rashmitha, Rao, B., Dsouza, R. A., Priyanka, & Acharya, R. P. (2024). IoT Based Infant and Environment Monitoring System. *2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON)*, 1–6. https://doi.org/10.1109/NMITCON62075.2024.10699170

Dumpalwar, D. (2024). IoT-Based Baby Cradle System. International Journal for Research in Applied Science and *Engineering Technology*, 12(8), 513–519. https://doi.org/10.22214/ijraset.2024.62869

Duthie, N., Ansari, K., Wang, M., & Ray, P. K. (2024). Design of a shared context-aware assistive robot for personalised aged-care services. *Digital Medicine*, 10(1). https://doi.org/10.1097/DM-2023-00017

Faulkner, M., Gerstenblatt, P., Lee, A., Vallejo, V., & Travis, D. (2016). Childcare providers: Work stress and personal well-being. *Journal of Early Childhood Research*, 14(3), 280–293. https://doi.org/10.1177/1476718X14552871

Gongor, F., & Tutsoy, O. (2024). On the Remarkable Advancement of Assistive Robotics in Human-Robot Interaction-Based Health-Care Applications: An Exploratory Overview of the Literature. *International Journal of Human–Computer Interaction*, 1–41. https://doi.org/10.1080/10447318.2024.2342101

Gupta, P. K., Varadharajan, N., Ajith, K., Singh, T., & Patra, P. (2024). Facial Emotion Recognition Using Computer Vision Techniques. *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1–7. https://doi.org/10.1109/ICCCNT61001.2024.10725699

Hsu, L.-J., & Šabanović, S. (2024). Designing AI for Partnership in Care. *Proceedings of the AAAI Symposium Series*, 4(1), 158–162. https://doi.org/10.1609/aaaiss.v4i1.31787

Joshi, P., & Shukla, S. (2019). Development in Early Childhood. In Child Development and Education in the Twenty-First Century (pp. 41–61). *Springer Nature Singapore*. https://doi.org/10.1007/978-981-13-9258-0_3

Journal, I. (2023). Smart Wearable Device for Child Safety by Using IOT. *International Journal of Scientific Research in Engineering and Management*, 07(12), 1–13. https://doi.org/10.55041/IJSREM27756

Lam, S., Williams, J., McBride, A., & Kelly, M. (2021). Putting Our Kids First: The Need for Systemic Changes in American Child Care. *Journal of the American Academy of Child & Adolescent Psychiatry* (*JAACAP) Connect*, 7(4). https://doi.org/10.62414/001c.92351

Li, F., Cui, C., & Hu, Y. (2024). Classification of Infant Crying Sounds Using SE-ResNet-Transformer. *Sensors*, 24(20), 6575. https://doi.org/10.3390/s24206575

Luhr, S., Schneider, D., & Harknett, K. (2022). Parenting Without Predictability: Precarious Schedules, Parental Strain, and Work-Life Conflict. RSF: The Russell Sage Foundation *Journal of the Social Sciences*, 8(5), 24–44. https://doi.org/10.7758/RSF.2022.8.5.02

Mishra, P., Verma, A. S., Chaudhary, P., & Dutta, A. (2024). Emotion Recognition from Facial Expression Using Deep Learning Techniques. *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, 1–6. https://doi.org/10.1109/I2CT61223.2024.10543313

Mizuki, R., Kyuzen, M., Nishizawa, S., & Mori, S. (2020). Assessing the Association between Care Quality and Household Chore Rules in Residential Care Institutions in Japan. *International Journal of Child, Youth and Family Studies*, 11(4), 80–95. https://doi.org/10.18357/ijcyfs114202019939

Morgan, K. J. (2021). Promoting Quality and Equality through Early Education and Care. In Politics, Policy, and Public Options (pp. 117–133). *Cambridge University Press.* https://doi.org/10.1017/9781108767552.008

Mortejo, J. (2024). Balancing Work and Family Commitments: Adaptive Strategies of Dual Income Couples. *International Journal of Empirical Research Methods*, 2(1), 72–78. https://doi.org/10.59762/ijerm20527579212024030517 0520

Nivedha, k, Anitha, K., & Jayaprakesh, D. (2020). Design and Implementation of Smart Day Care Monitoring System. *European Journal of Molecular & Clinical Medicine*, 7(09), 1553–1557. https://doi.org/10.31838/ejmcm.07.09.169

Pathak, A., Ogunbayo, A., Trushna, T., Khare, S., Mathur, A., Atkins, S., & Diwan, V. (2022). Perceptions and Experiences of Caregivers on Child Injuries: A Qualitative Study from Central India. *Journal of Prevention*, 43(4), 549–565. https://doi.org/10.1007/s10935-022-00682-3

Vol 1(2), DOI:10.82036/tr28me67, July 2025 ISSN:2977-9332

Phillips, D. A., & Lowenstein, A. E. (2011). Early Care, Education, and Child Development. *Annual Review of Psychology*, 62(1), 483–500. https://doi.org/10.1146/annurev.psych.031809.130707

Pramod, K., Katkar, A., Kanukuntla, M., Bodakunta, J., Yarraboina, K., & Jamalapuri, S. (2024). Smart Device for Child Safety with Parental Alerts. 2024 *2nd International Conference on Intelligent Data Communication Technologies and Internet of Things* (IDCIoT), 1617–1623. https://doi.org/10.1109/IDCIoT59759.2024.10467637

Roy, S. (2020). Faulty Traditional Practices in Infant Care. *International Journal of Medical and Surgical Sciences*. https://doi.org/10.32457/ijmss.v7i2.555

S Raju, M., & T, V. (2022). IoT Based Artificial Intelligence Controlled Environment Monitoring Robotic System. *International Journal of Science and Research (IJSR)*, 11(7), 942–945. https://doi.org/10.21275/SR22714001717

Scheibling, C., Young, M., Milkie, M. A., & Schieman, S. (2023). Gender, work, and the family's morning rush hour: the strain associated with preparing children for the day. Community, *Work & Family*, 1–18. https://doi.org/10.1080/13668803.2023.2298313

Seiler, C. W., Sticca, F., Gasser-Haas, O., & Simoni, H. (2022). Long-Term Promotive and Protective Effects of Early Childcare Quality on the Social–Emotional Development in Children. *Frontiers in Psychology*, 13. https://doi.org/10.3389/fpsyg.2022.854756

Shabunova, A., & Leonidova, G. (2023). Work-family balance: assessments of successful working parents. *Population*, 26(1), 123–134. https://doi.org/10.19181/population.2023.26.1.10

Shameema, L. R., & L, S. (2023). Wearable Smart Gadget for Child Monitoring based on the Internet of Things. *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 1827–1831. https://doi.org/10.1109/ICACCS57279.2023.10113009

Shen, X., Lv, B., Liu, T., & Cheng, Q. (2024). Infant Speech Emotion Recognition Based on Channel Attention Mechanism with ResNet-BiLSTM. *2024 5th International Conference on Information Science, Parallel and Distributed Systems (ISPDS)*, 54–57. https://doi.org/10.1109/ISPDS62779.2024.10667573

Singh, G., Shekhar, A. R., Yu, X., & Saniie, J. (2023). Smart Infant Monitoring System Using Computer Vision and AI. *2023 IEEE International Conference on Electro Information Technology (EIT)*, 1–6. https://doi.org/10.1109/eIT57321.2023.10187295

Sumathi, V., Mohamed Abdullah, J., Senthil, R., & Prema, E. (2023). Smart Child Tracking System. In Smart Grids for Smart Cities Volume 2 (pp. 1–9). *Wiley*. https://doi.org/10.1002/9781394216796.ch21

Sung, Y., Chen, Z., Das, J., & Tokekar, P. (2023). A Survey of Decision-Theoretic Approaches for Robotic Environmental Monitoring. Now Publishers. https://doi.org/10.1561/9781638283058

Yinusa, M. M. (2024). Comprehensive Analysis of Family-Based Childcare Options for Orphaned Children. *Newport International Journal of Current Issues in Arts and Management*, 4(1), 8–11. https://doi.org/10.59298/NIJCIAM/2024/4.1.8113