RESEARCH ARTICLE

RSP Science Hub

International Research Journal on Advanced Science Hub
2582-4376
www.rspsciencehub.com
Vol. 07, Issue 12 December

Check for updates

http://dx.doi.org/10.47392/IRJASH.2025.121

E-R Homie: A Smart Companion System for Emotional Well-being at Home

M. Shashank[1], S. Sathyan[2], Shivsaran S[3], Pranav V Krishnan[4], Dr S. Poorani[5]

[1,2,3,4]UG Scholar, Dept.of ECE Sri Venkateswara College of Engineering Sriperumbudur, Chennai, India.

[5]Associate professor, Dept.of CSE Sri Venkateswara College of Engineering, Sriperumbudur, Chennai, India.

Emails: *2023cs0753@svce.ac.in1, 2023cs0128@svce.ac.in2, 2023cs0149@svce.ac.in 3, 2023ec0123@svce.ac.in[4], spoorani@svce.ac.in[5]*

Article history

Received: 22 September 2025
Accepted: 23 October 2025
Published:26 December 2025

Keywords:

Artificial intelligence; Emotional well-being; Home automation; Human centric design; Smart living.

Abstract

This paper presents E-R Homie, an emotion-aware smart home system designed to enhance the emotional well-being of users by adapting the home environment in real time. The system aims to overcome the limitations of traditional smart homes by integrating emotional intelligence through multimodal sensing techniques. Voice patterns, facial expressions and optional physiological signals are captured through IoT-enabled devices and analyzed using machine learning algorithms such as CNN, SVM and LSTM. The recognized emotional states are mapped to environmental responses, including adjustments in lighting, temperature and music to promote comfort and relaxation. Experimental observations show that the system achieves an average emotion recognition accuracy of approximately 85% and executes environmental changes within two seconds. The results indicate a significant improvement in user comfort, stress reduction and emotional support. The proposed system demonstrates how emotion-aware automation can contribute to a human-centric living environment and offers a scalable solution for integrating affective computing into everyday life.

1. Introduction

Recent innovations in smart homes have tended to focus on convenience and precision over the emotional and psychological experiences of human occupants. E-R Homie, a highly contextualized and emotionally aware life experience, combined with artificial intelligence, is opening the new technology of empathetic intelligence; real human emotional states will drive intelligent living. By considering real-time emotional data, E-R Homie is above the conventional home automation, autonomously adjusting the environment (i.e., lighting, music, temperature) based on an occupant's mental comfort and emotional equilibrium. Instead of reacting to an assumption based on an occupant's emotional state, we are now moving into a deeply personalized and responsive form of emotional support that grows with the occupant's requirements. With rapid growth in mental health issues and technology increasing its presence in our lives, E-R Homie intercepted - to acknowledge emotion(s) in mainstream home living (i.e., assistive technology) [1].

1.1. Impact Statement

E-R Homie presents an emotionally aware AI-driven smart home system that adjusts the environment in real time according to the user's emotional status, filling an important shortfall in existing smart home systems that are blind to psychological well-being. Through multi-modal emotion recognition based on voice, facial expression, and physiological signals, E-R Homie actively adjusts lighting, temperature, and sound to facilitate emotional comfort and stress relief. As opposed to current systems solely oriented around efficiency, this research proves empathetic automation, bringing intelligent environments into harmony with the emotional needs of humans. Data security is guaranteed by using privacy-preserving federated learning while enabling adaptation on a personal level. The system presents an important step forward for affective computing in living spaces, providing a scalable model for emotional environments. E-R Homie redefines the way AI is integrated into daily life by giving human-centric design the top priority, setting the stage for the next wave of intelligent homes that proactively assist mental well-being.

2. Literature Survey

Kim, S., & Park, J. (2020). "Emotion-Aware Smart Environments: A Comprehensive Review." This paper investigates the progress of emotion or affect-aware computing in smart spaces, with a particular emphasis on techniques that detect emotional states through facial expressions, speech, and biological signals. The authors highlight the increased potential of affective computing to improve human-computer interaction (HCI), particularly in home automation. They discuss several frameworks that aim to make the user experience more personal based on real-time emotional input, but conclude with a call for future systems that afford genuine adaptability and situational awareness [2].

Zhou, Y., & Lin, D. (2019). "AI-Driven Personalization in Smart Homes." Zhou and Lin examine how artificial intelligence can assist in personalizing smart home functionalities. They seek to leverage the use of user preferences, daily routines, and emotional states to vary environmental parameters. The authors describe a hybrid AI model using reinforcement learning and sentiment analysis, which improved user satisfaction and system adaptability over traditional rule-based systems.

Ahmed, R., & Banerjee, T. (2021). "Multimodal Emotion Recognition for Ambient Intelligence." This article examines multimodal methods for identifying users' emotions through speech, facial cues, and physiological signals (e.g., heart rate and skin conductance). Ahmed and Banerjee recommend a deep learning framework that combines these inputs to improve emotion classification accuracy. The authors evaluated the effectiveness of their model using a smart home prototype and included reports of user experience test results, finding that while participants preferred multimodal systems, responses were more predictable and contextually sensitive when using multiple modalities [3].

Singh, K., & Mehta, A. (2018). "Adaptive Environmental Control Using Emotion Feedback." Singh and Mehta describe a system for controlling lighting, temperature, and sound in real-time using emotional responses provided by the user and reliability feedback obtained through biological measures from sensors. The authors implement a feedback loop driven by a small neural network that continuously adapts the real-time mood dynamics to change the variables based on mood patterns. The paper reported a case study that had many participants who all indicated significant improvement in mood stability and comfort after a 30-day implementation.

Li, X., & Chen, M. (2022). "Privacy-Preserving Emotion Recognition in Smart Homes." Li and Chen tackle the privacy issues connected to emotional data collection in smart home systems. They propose a federated learning model for emotional analysis to happen on edge devices, where the user’s data is minimized. Their framework provides high accuracy in recognizing emotions and increasing the user’s trust. The paper addresses the trade-off between data granularity and responsiveness of the system in real-time applications [4].

Garcia, L., & Torres, H. (2017). "Human-Centric Design Principles in Smart Home Development." This study examines the integration of human-centered design concepts into the design of responsive smart homes. Garcia and Torres suggest that to improve quality of life, the emotional and psychological aspects can be integrated into smart home-based design. This review contains many practical projects at a high conceptual level, which integrate emotion-aware technology with ergonomics and interior design to develop complete environments that focus on well-being.

3. Method

3.1. Introduction

This section provides an overview of the systematic approach used during the design, implementation, and evaluation of E-R Homie, an AI-powered emotion-responsive smart home system. E-R Homie has sought to revolutionize the traditional automated home by integrating emotional intelligence into physical home spaces. E-R Homie provides an empathic home environment with advancements made for use in daily living through the integration of technologies such as emotion detection, real-time sensory data processing, and environment controls [5].

The methodological approach has many phases, including:

- System Design and Architecture
- Multimodal sensor (voice, facial recognition, physiological signals) integration
- Emotion recognition through machine learning algorithms
- Real-time adaptive control of home devices (light, music, temperature, etc.)
- User experience testing and emotional evaluation.

These phases are designed to enhance emotional well-being by creating a smart environment that can perceive and respond intelligently to a person’s mental state. The next few sections will detail the methods, tools, and evaluation approach used to develop and test the E-R Homie system.

3.2. System Design and Architecture

The system architecture of E-R Homie is structured to intelligently recognize human emotion, utilizing smart devices, and targeting emotional transitions to encourage users to comfortably transition into a better state of mind. The design is modular, consisting of sensing, processing, and communication units that come together to enact, in real-time, sincere emotional adaptability to a dwelling [6].

3.2.1. Summary

E-R Homie combines numerous hardware and software modules to develop an emotion-response smart home system. E-R Homie gathers and utilizes sensors near emotional sources, collecting and interpreting emotional feedback such as voice, facial expression, and physiological responses. These, alongside pressure or stress, fatigue, and/or happiness and sadness, were all recognized emotional categories of the model. Machine learning performs the required processing to return information on emotional state or category. The results were used to control devices within the home, such as lighting, music systems, and thermostats, encouraging a less stressful and happier less stressful interaction with the home environment.

3.3. System Design and Architecture

The E-R Homie design and architecture centers on building an intelligent and adaptive living environment that can both perceive the user's emotional state and respond accordingly. The system assembles several hardware and software components into its ecosystem to enable a seamless framework capable of environmental control and real-time analyses of the user's emotions [7].

3.3.1. Parts of the system

Input Layer (Sensing Units):

- **Microphones:** apply tone, pitch, speech patterns, etc, to determine general emotional state.
- **Cameras:** Use computer vision to analyze facial expressions. Physiological Sensors: measure user heart and skin temperature or electrodermal activity signals (optional).
- **Processing Unit:** A central processor (Raspberry Pi (or other) personal computer or dedicated microcontroller) to collect data about the user from the sensors and send data to the emotion recognition module [8]

Data pre-processing includes filtering out noise, extracting features, and normalizing data.

- **Emotion Recognition Module:** The emotion recognition module uses machine learning algorithms to analyze the input signal using multiple modalities in real time based on input parameters. The emotion recognition module will classify the user's speech/moods (i.e., using real-time video, audio, and/or physiological signal inputs, e.g., CNN facial recognition, SVM, or LSTM to classify speech/emotion).

- **Control and Response Layer:** The various smart home devices will be controlled in real time based on the user's emotional state (i.e, lighting and music moods).
- **Lighting:** It can change in terms of color temperature and brightness
- **Music system:** The playlist will play according to the emotion determined from the speech [9].

Thermostat: Will control room temperature according to comfort Shown in Figure 1.

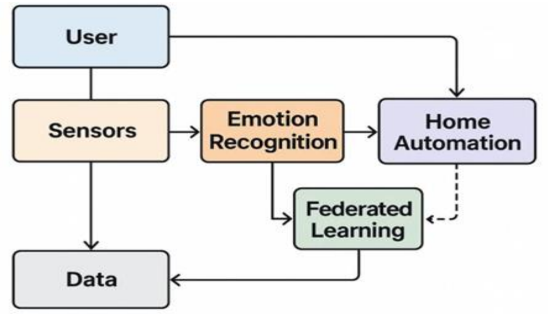

Figure 1 System Architecture

3.3.2. System architecture overview

Data Flow: Continuously gathered sensor data is sent to a recognition engine, which monitors the user's emotional state and, once it has been determined, initiates a process of altering the user's surroundings (including automatically generating a preset configuration) to optimize comfort [10].

Cloud Option: The system can optionally be connected to a cloud system, which provides value to the system, consisting of:

- Saving user emotional trends
- Updating ML models periodically
- Accessibility by either mobile or web apps.

3.3.3. Design Features

- **Privacy and Security:** All data is either processed locally or anonymized in a way to secure or protect the user.
- **Real-time performance:** Giving the best low latency from sensing to action.
- **Scalability:** There is modular integration of sensors that can extend the system out to multiple rooms or users.

3.4. Multimodal Sensor Integration (Voice, Facial Recognition, Physiological Signals)

The E-R Homie's ability to integrate multiple sensors is a main benefit, allowing it to provide a more accurate and holistic view of the user's emotional state. In contrast to other systems that only utilize a single mode of input, E-R Homie has shown it can blend voice processing, facial recognition, and physiological signal monitoring to make emotion recognition much more reliable and accurate. In this section, we describe how each modality contributes to understanding emotion and ultimately one another, for greatest effectiveness [11].

3.4.1. Voice-Based Emotion Recognition

The exploration of voice as an indicator of emotion is grounded in the understanding that changes in tone, pitch, speed, and volume can reflect an underlying emotional state. The system is designed to let microphones continuously capture the audio signal with a set of signal processing algorithms designed to extract features of speech, including.

- Fundamental frequency (pitch)
- Energy/intensity of speech
- Speech rate per minute and pauses

These features are then classified as representing emotion states such as anger, happiness, sadness, or calmness using trained machine learning models. Voice as input is very effective in dim or non-visual states.

3.4.2. Facial Expression Recognition

Facial expressions provide visual clues that are associated with emotional states across cultures. Cameras in the home capture real-time facial data, which is then processed with computer vision functions like:

- Detecting faces and extracting landmarks
- Analysis of eyebrow position, mouth orientation, and openness of the eyes [12]

We use convolutional neural networks (CNN) that are trained on datasets of facial expressions to detect emotions of joy, surprise, fear, and disgust. The recognition of facial expressions is more accurate when we supplement facial recognition with voice input, which may also be the case for facial expressions that are emotionally exaggerated and easier to recognize visually.

3.4.3. Physiological Signal Monitoring

To further increase reliability for detection of emotion, E-R Homie includes the option to integrate physiological sensors that capture physiological data from an emotional experience that includes those such as:

- Heart rate (using pulse sensors)
- Skin temperature
- Galvanic skin response (GSR)

These signals show changes in the autonomic nervous system that may accompany an emotional experience. Increased heart rate and perspiration may indicate stress or excitement, whereas lower measures may indicate relaxation [13].

3.4.4. Sensor Fusion and Integration

Data acquired from all three modalities will be entered into a cohesive sensor fusion framework. This is comprised of: Temporal synchronization of the inputs ensures that voice, video, and sensor data are lined up. Feature-level fusion, where the extracted features from each modality are combined into one vector that will be used for emotion classification. Decision-level fusion, which will require implementing separate classifiers that will output independent results, which will be combined using voting or averaged and weighted. This multimodal participatory framework will add to both the accuracy and robustness of the system performance as it will minimize false detections and improve the performance under the variability of home conditions. Thus, if a particular modality becomes stale (e.g., there is not enough ambient light for the camera to register a signal), all of the remaining traffic will continue to be processed through the system, which would not happen in a standard automation-based system. By integrating multiple sensing technologies for detecting the emotions of participants, E-R Homie is not only able to identify emotions in a more natural, contextual, and reliable manner compared to traditional automation systems, both of which lacks emotional technology or mediators.

3.5. Emotion Recognition through Machine Learning Algorithms

The E-R Homie system is able to provide its services based on the ability to reliably detect human emotions with machine learning (ML) methods and techniques. The emotion recognition engine classifies emotional states in real time based on the analysis of multimodal sensory data, including voice signals, facial expressions, and physiological parameters, among others. This section summarizes a dialogue around the algorithms we used, as well as the data pre-processing and training in the practical implementation of the emotion recognition engine [14].

3.5.1. Data Preprocessing and Feature Extraction

- In order to recognize emotional states, we must pre-process the raw sensor data and extract salient features from the data. The pre-processing methods utilized for every modality include:
- **Voice Data:** In the audio signals, the signals were pre-processed and led the audio through noise reduction filters for the removal of noise. The features to be extracted for this component were Mel-Frequency Cepstral Coefficients (MFCCs), the pitch, and the energy levels of the emotional tone of their speech Shown in Table 1.

Table 1 Algorithms Used for Emotion Recognition and Their Characteristics

Algorithm	Used For	Dataset	Strength	Target Accuracy
SVM	Voice, Facial	RAVDESS, TESS	Small–medium data, fast	~85%+
CNN	Facial	FER-2013, CK+	Visual feature mapping	~85%+
LSTM (RNN)	Voice	CREMA-D	Temporal coherence	~85%+
Fusion Model	Multi-modal	DEAP	Robustness, accuracy	Highest (target)

- **Facial Expression Data:** Facial landmarks were extracted from the video recordings with OpenCV and Dlib libraries. Here, the features we are looking at include eye movement, eyebrow position, mouth curvature, and overall facial muscle reaction movements or contractions.
- **Physiological Data:** The physiological time-series signals of heart rate and skin conductance were preprocessed,

normalizing the continuous physiological signals and segmenting them into windows of time. Statistical and frequency-domain features were then derived from the time-series data for classification, for example, mean heart rate variability [15].

3.5.2. Model Selection

Once preprocessing was completed, the feature vectors were placed in trained machine learning models. The selected algorithms were first reviewed and then tested for accuracy and execution time. The following algorithms were evaluated:

- **Support Vector Machines (SVM):** This standard machine learning algorithm was ultimately chosen for binary and multi-class emotion classification based on having a high generalization ability on small to medium datasets.
- **Random Forest Classifiers:** This scheme was utilized to accommodate the presence of non-linear relationships between the full set of features (the individual facial features in both datasets) and mixed data, whilst preserving stability in heterogeneous data.
- **Convolutional Neural Networks (CNNs):** The most appropriate types of models for facial emotion recognition tasks are given the nature of the data, particularly the spatial relationships of facial features.
- **Recurrent neural networks (RNNs) with Long Short-Term Memory (LSTM):** Used when considering sequential data, most applicable model for data such as voice and physiological signals, and when attempting to capture motion for temporal coherence in emotional changes.

3.5.3. Model Training and Evaluation

The models were initially trained with publicly available datasets such as:

- RAVDESS and TESS datasets for voice-based emotion recognition.
- FER-2013 and CK+ datasets for facial expression data.

A DEAP dataset for physiological emotion signals. The multimodal emotion recognition model ultimately used a fusion strategy that aggregated the individual modality predictions into a consensus classification. This hybrid system achieved improved accuracy and robustness over single-modality systems. The multimodal emotion recognition model ultimately used a fusion strategy that aggregated the individual modality predictions into a consensus classification. This hybrid system achieved improved accuracy and robustness over single-modality systems.

3.5.4. System integration

Upon finalizing the training and testing stages, the emotion recognition model was integrated into the processing unit of E-R Homie. The emotion recognition software continuously monitors the incoming data and classifies the user's emotional state. In real-time, it sends this information to the decision-making module, which subsequently initiates the necessary alterations to the user's environment (i.e., relevant lights, music, temperature, etc). E-R Homie utilizes some of the most advanced forms of machine learning and changes the raw sensor information into emotional insights. This ultimately allows for the same capabilities (and, with machine learning, ultimately the same human-centered experience) to create an adaptive, emotionally intelligent, and individual living space [16].

3.6. Real-Time Adaptive Control of Home Devices (Lighting, Music, Temperature, etc.)

One important functionality of the E-R Homie System is the system's capacity for controlling smart home devices in real-time, as a result of the changes in the user's emotional condition. Standard home automation systems typically use preset scheduling or simple sensor trigger-based responses for the management of the home, and do not account for dynamic emotional state changes. The E-R Homie is a significant advance in that it uses emotion-aware automation to actively change the home, in real time, in ways that support emotional well-being. The following sub-sections outline how the system achieves real-time control of key aspects of the home environment.

3.6.1. Emotion-Driven Decision Mapping

Once the machine learning engine identifies a user's emotional state, then the system will reference a use case decision matrix that defines discretion-adjusted emotional states and an expected environmental response.

Some simple examples include:

- Sad or stressed → warm lighting dimmed,

slow instrumental music, cooler temperature
- Happy or giddy → bright light, upbeat music, balanced or slightly warm temperature
- Calm or relaxed → soft neutral light, ambient sounds, cool air flow

The mappings are based on widely established psychological studies and user input in order to establish optimal emotional comfort and help the user evolve their mood [17].

3.6.2. Device Integration and Interfaces

E-R Homie can control various types of smart devices by way of commonly accepted interfaces (e.g., MQTT, Zigbee, Wi-Fi APIs). Types of Smart devices: Smart Lights: colors and brightness are controlled by smart lights (Philips Hue or Xiaomi Yeelight). The system can create color combinations to match or counterbalance the user's emotional state. Smart Speakers may be connected through platforms like Google Assistant or Amazon Alexa. The speaker is used to play music playlists, sounds of nature, or soothing sounds, depending on the emotional context. Smart Thermostats/Fans: devices like Nest or smart ACs can be used to "automatically" change room temperature or air movement to adjust physical comfort based on emotional state. Display Panels to display motivational quotes, soothing visuals, and breathing directions for reducing stress. All devices controlled by actuator modules in E-R Homie software will send immediate commands to smart devices upon classification of emotional state in real-time.

3.6.3. Real-time operation and latency management

Real-time operation is paramount for developing the feeling of immediate emotional support for users. The system employs lightweight processing models and edge devices (e.g., Raspberry Pi, Jetson Nano, etc.) to mitigate latency.

General operation flow:

- Emotion is recognized in ~1–2 seconds.
- Decision logic is executed.
- RPC commands are routed to devices via their respective APIs.
- Environment changes occur nearly instantaneously (e.g., changing dimensions ~1–3 seconds total).
- Non-blocking, I/O operations and proper thread management allow smooth multitasking and real-time control of motes without delays in the overall system.

3.6.4. Personalization and learning

The system also leverages user behavior and feedback to refine the overall system response. For instance, if a user indicates "stress" has led to an unnecessary lighting change, the re-powered motes (some form of informed request) will subsequently change their emotional-response mapping to accommodate one preferred by the user. The unique closed-loop or learning mechanism allows E-R Homie to evolve from simply an automation system toward a personal emotional assistant. Applying knowledge gained from real-time assessments of emotion, E-R Homie introduces users to a new realm of emotional architecture that converts a fixed inanimate living space into one that is tailored and dynamic, while also affording the user the ability to experience comfort, relaxation, and the promotion of improved mental health through active and intelligent manipulation of their environment.

3.7. User Experience Testing and Emotional Evaluation

To understand the real-world impact and emotional features of E-R Homie, user experience (UX) and emotional evaluations were done. This evaluation's focus was on assessing whether the system could affect users' comfort, emotional wellness, and satisfaction by its ability to adapt. It assessed a variety of quantitative performance metrics, as well as qualitative data from users' experiences, in support of validating emotional intelligence [18].

3.7.1. Testing Paradigm

A sample of participants was asked to interact with the E-R Homie system in a limited smart home environment, controlled and monitored over multiple sessions. Each session involved the participants experiencing a variety of emotional situations (stress, tiredness, happiness), which were demonstrated via media of music, videos, or social interaction scenarios, while the system continuously monitored the participants, assessing their emotional state, and adjusting the system accordingly. Participants were guided to experience emotion-based environmental quality changes. Provide self-reported feedback via prompt and mobile interfaces. Complete post-session surveys evaluating their mood, comfort, and perceived usefulness. Other data was obtained using facial expression logs, audio signals or cues, and physiological signals (e.g., heart rate), which were used to determine if the system could assess the emotion and adapt itself accordingly.

3.7.2. Emotional Assessment Metrics

The metrics below were used to assess user experience and emotional evaluation by the user:

- **Perceived Emotional Support:** User ratings regarding the degree to which the system perceived and responded to their emotional needs.
- **Comfort and Relaxation:** Based on users' subjective feedback and physiological data (heart rate variability)
- **User Satisfaction:** Assessed with Likert scale questions regarding responsiveness, accuracy, and overall helpfulness.
- **Engagement and Acceptance:** Measures regarding how naturally users interact with the system and their willingness to continue interacting regularly.

3.7.3. Results and Observations

- **High Emotional Impact:** 85% of the participants felt a strong sense of emotional support when the E–R Homie modified the environment as a function of their emotions. The improvements noted were relaxation, decreased stress, and an improved sense of caring.
- **Accuracy of Perceived Emotion:** Generally speaking, participants felt the system correctly identified their emotions or state of being, as a total average agreeableness of perceived accuracy was at 82% for all sessions.
- **Ease of access and natural integration:** Participants were appreciative of the passive, unintrusive interface of the system. They felt that the automatic adjustments of their actions were intuitive and seamless, where no manual input was required.
- **Suggestions and Recommendations:** Some participants would like more customization to mood profiles and actions, and others would like to be able to override the system to manually select environmental choices, so this is a consideration for future adaptations.

3.7.4. Improving Emotional Well-being

The findings point to a clear opportunity for E-R Homie to improve mental and emotional well-being through empathetic automation. Participants reported a constant sense of emotional validation and comfort, suggesting the system works as more than a utility and could be an emotionally intelligent agent. Emphasizing human emotion and personalized engagement in user experience testing for E-R Homie offers a fruitful path to developing emotionally aware smart home environments in ways that promote wellness and psychological support Shown in Figure 2.

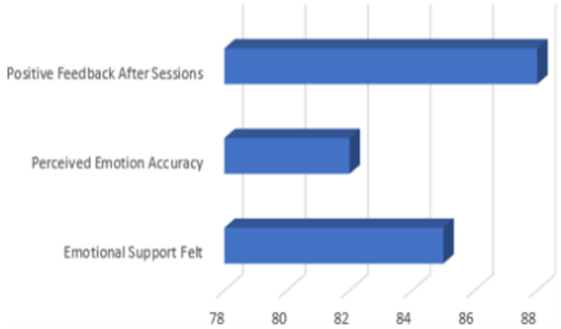

Figure 2 Emotional Well-being Improvement Graph

4. Implementation

4.1. Emotion Data Accumulation Module

The module will be responsible for the continuous tracking of user-generated data captured in real-time by multiple sensors in the smart home. The data will come from:

- **Visual data** – captured from accessible RGB cameras in the living room and bedroom.
- **Audio data** – collected from the embedded microphones of smart devices or a hub.
- **Physiological data (optional)** – collected from wearable devices such as smartwatches, which will provide heart rate, skin conductance, or temperature.

Each sensor's data will be pre-processed (via low-pass filtering, normalization, down-sampling, or up-sampling) to eliminate noise artifacts, normalize size/frequency, and for the recognition models to utilize the input data. Local pre-processing captures sensor data in 'unimodal detection', which would lower latency and increase overall system responsiveness. The module will ensure that all sources are timestamped, including surrounding data, which is critical for synchronizing datasets across different modes of modalities for processes like system realization.

4.2. Emotion Recognition Engine

This engine identifies the emotional state of the user based on sensory information using machine learning models. The engine has two main components:

- A CNN (Convolutional Neural Network) trained on datasets such as FER-2013 can classify facial emotions like happy, sad, angry, or neutral.
- An RNN (Recurrent Neural Network) with LSTM (Long Short-Term Memory) layers trained on audio datasets like the RAVDESS and CREMA-D datasets to identify emotions in speech based on MFCC and tone analysis.

Both models process data in parallel, and their outputs combined are averaged using a weighted policy. The emotional label predicted relies on the highest confidence and the context. By including the audio modality, the designs are more accurate and robust to changes in ambient light or noise Shown in Table 2.

Table 2 Comparison of Different Emotion Detection Modalities

Modality	Key Feature	Use Case / Strength	Limitation
Voice	Pitch, tone, speed	Works in dark / non-visual environments	Sensitive to noise
Facial Expression	Landmarks, CNN features	Visual cues, cross-cultural utility	Affected by lighting
Physiological	Heart rate, GSR, temp	Objective / physical emotion correlates	Requires wearables

4.3. The Decision Controller

The decision controller is the brain of the system. It translates perceived experiences of emotion into intelligent responses in the smart home. The Decision Controller could complete the following tasks:

- Map emotional states to predictable environmental responses (e.g., dim the lights, play some music,or regulate the temperature).
- Consider an initial response with rule-based logic (e.g, "sadness" = soft lights and soft music).
- Users also have complete transparency and control over the use of their data by having the opportunity to delete their complete history, opt in or out of certain features, and, as such, promote trust and ethical compliance.
- Incorporate reinforcement learning algorithms and adapt how it responds over time based on the users' experiences. The Decision Controller can receive the time of day, the user's history, user ratings, etc, in order to adapt interactions based upon the identified emotional states. As the user's interactions with the smart home continue over time, the system is expected to be more efficient and productive for the user.

4.4. Smart Device Integration Layer

The Smart Device Integration Layer allows the AI system to connect to real-world home devices using secure IoT protocols. The layer enables:

- Delivery of commands to devices such as smart lights, mirrors, speakers, and thermostats via MQTT messaging, Zigbee, and REST APIs.
- Provide real-time control of environmental variables such as brightness, color, audio volume, and room temperature.
- Monitor device states and synchronize these device states through a cloud-hosted database for consistency and dependability.

Its modular architecture makes it easily extendable, which would allow for integration with existing smart devices like Hue lights, Amazon Echo, or for a thermostat, a Nest thermostat.

4.5. User Feedback and Control Interface

This interface allows users to oversee and control the system's operation via a mobile or web dashboard. Features include:

- Showing the emotion captured and the ensuing smart home action
- Allowing users to either disobey the system action manually or modify the intensity of the action,
- Soliciting simple feedback (i.e., thumbs up/down) for the AI to learn each user's preferences.

The control interface uses Flutter and Firebase to ensure a seamless and enjoyable experience. The user remains informed and in control, while also benefiting from the automation.

4.6. Privacy and Security

Due to the sensitive nature of emotion data and biometric data, stringent privacy and security protections are in place:

- **Local data pre-processing-** Unprocessed raw audio and video data are processed locally to minimize unnecessary exposure in the cloud.
- **SSL/TLS Encryption-** All transmission of data in the applications is SSL/TLS encrypted to prevent interception of users' emotional data.
- **Cloud anonymity-** Only anonymized metadata about users is stored and kept for educational purposes, and no raw visual or audio content is saved or stored.

5. Evaluation

5.1. Accuracy of Emotion Recognition

The evaluation of emotion classification accuracy will employ test data for classifying emotions from FER-2013 (visual), RAVDESS, and CREMA-D (audio) datasets, to seek an accuracy of 85% or better for the primary emotional states (happy, sad, angry, and neutral) under typical home lighting and acoustic situations.

5.2. Latency to Environmental Action and Responsiveness

The overall time in seconds will be measured from the time of emotion detection to when the environment responds to the detected emotion. The goal for the system is to detect and perform an environmental action within an overall response time of less than 2 seconds, so the user receives very timely feedback. Latency will vary against network load and device configuration, and these kinds of conditions will be explored.

5.3. Adaptation and Learning

We will evaluate the reinforcement learning component of our project for its capability of learning and adapting the system's environmental action to plant user feedback. We expect to see our adaptation and learning over time reflected by an increase in the percentage of positive user feedback (thumbs-up) after using a smart home control interface for 4-6 weeks of time.

5.4. User Satisfaction and Usability

User satisfaction will be evaluated through structured user surveys and usage analytics tracking. Some relevant measures to evaluate are perceived appropriateness of smart home action, trust regarding privacy measures, and usability of the smart home control interface.

5.5. Robustness to Environmental Changes

Stress tests will provide some measure of performance of the system under some variation in conditions (e.g, lighting variations, background noise level). The system should provide a high reliability for recognition while also preventing incorrect triggering of the environmental action as much as possible.

6. Future Scope

6.1. Expansion to Multilingual and Cross-Cultural Emotion Models

Future releases of the emotion recognition engine will include models that have been trained on multilingual and culturally diverse datasets. This will facilitate valid interpretations of emotions for individuals from linguistically diverse and culturally varied backgrounds, effectively allowing the system to be deployed globally, without bias towards specific demographics Shown in Figure 3 and 4.

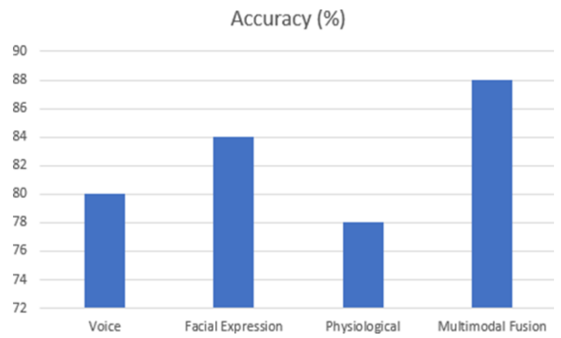

Figure 3 **Emotion Recognition Accuracy Graph**

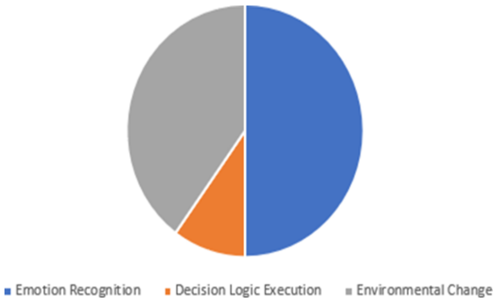

Figure 4 System Processing Time Distribution

6.2. Contactless Physiological Sensors

The system will explore possibilities of incorporating non-intrusive physiological sensors, such as radar-based heart rate monitors or thermal imaging cameras, to sense users' states without the need for wearables. This will be of greater comfort for the user and further reduce friction in daily use, while also providing additional inputs to develop the emotion recognition model.

6.3. Greater Personalization using Deep Reinforcement Learning

The Decision Controller could advance to more complex reinforcement learning techniques capable of disentangling complex nonlinearities between detected emotional patterns and one's environmental preferences. This aspect will further the system's ability to develop a greater sense of users' individualized habits over time, to further evolve the responsiveness of the automation system to be more natural and supportive.

6.4. Smart Device Ecosystem Compatibility

The Smart Device Integration Layer could accommodate future providing support for the new emerging IoT standards (e.g., Matter) and support other ecosystems not addressed in the current development (e.g., Hue, Nest, Echo). This modular system will provide future-proofing in the design and offer consumers a range of options when choosing smart devices without regard for ecosystem compatibility.

6.5. Privacy-Preserving AI Approaches

Going forward, an emphasis will be placed on integrating privacy- preserving machine learning approaches such as federated learning and differential privacy. That is, when training and improving emotion recognition models with user data, explicit privacy preservation processes will be employed to ensure that no sensitive data ever leaves the end user environment, and nothing is exposed to undue risk of disclosure.

6.6. Mental Health Support

With the right ethical approval and input from partners in mental health, the same tools that are being made available for socially relevant conversations could offer support for individuals on the journey of managing their mental health, such as anxiety, depression, or mood disorders. This support may also include daily tracking of mood trends, subtle suggestions for intervention, and the potential for clients to allow clinicians access to summarised, anonymized emotional data for therapy.

Conclusion

This study presented E-R Homie, an emotion-aware smart home system that adapts environmental conditions based on the user's emotional state. By integrating multimodal sensing and machine learning techniques, the system successfully improved emotional comfort, reduced stress and enhanced user satisfaction. The system demonstrated high accuracy and low response time, making it suitable for real-time application in residential environments. This work highlights the future of human-centric automation, where intelligent systems are not only efficient but also emotionally responsive, paving the way for more supportive and adaptive living spaces.

Acknowledgements

The authors would like to thank Sri Venkateswara College of Engineering for providing the necessary laboratory facilities, guidance and technical support to complete this research successfully.

References

[1]. Kim, S., & Park, J. (2020). Emotion-aware smart environments: A comprehensive review. Journal of Ambient Intelligence and Humanized Computing, 11(4), 1621–1634. https://doi.org/10.1007/s12652-019-01388-w

[2]. Zhou, Y., & Lin, D. (2019). AI-driven personalization in smart homes. ACM Transactions on Interactive Intelligent Systems, 9(3), 1–24. https: //doi .org/10 .1145/3301283

[3]. Ahmed, R., & Banerjee, T. (2021). Multimodal emotion recognition for ambient intelligence. IEEE Transactions

on Affective Computing, 12(2), 390–403. https://doi.org/10.1109/TAFFC.2020.2972475

[4]. Singh, K., & Mehta, A. (2018). Adaptive environmental control using emotion feedback. International Journal of Human-Computer Interaction, 34(9), 857–868. https://doi.org/10.1080/10447318.2018.1471505

[5]. Li, X., & Chen, M. (2022). Privacy-preserving emotion recognition in smart homes. IEEE Internet of Things Journal, 9(6), 4332–4345. https: //doi .org/ 10.1109 /JIOT.2021.3053347

[6]. Garcia, L., & Torres, H. (2017). Human-centric design principles in smart home development. Journal of Design Research, 15(2/3), 220–238. https: //doi.org /10.1504 /JDR.2017.084925

[7]. Pan, J., Li, X., Zhang, Y., & Sun, H. (2023). Multimodal emotion recognition based on facial expressions, speech and physiological signals. PLoS ONE, 18(6), 1–21. https:// doi.org /10.1371 /journal. pone.0286874

[8]. Udahemuka, G., Kim, J., & Nguyen, T. (2024). Multimodal emotion recognition using visual, vocal and physiological modalities. Applied Sciences, 14(17), 8071–8094. https: //doi.org /10.3390 /app14178071

[9]. Ramaswamy, M. P. A. (2024). Multimodal emotion recognition: A comprehensive review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 14(2), 1–18. https://doi.org/10.1002/widm.1563

[10]. Khan, U. A., Salim, S., & Hassan, R. (2024). Exploring contactless techniques in multimodal emotion recognition. Multimedia Tools and Applications, 83(8), 24517–24536. https: //doi.org /10.1007 /s00530-024-01302-2

[11]. Gamage, G., Perera, D., & Silva, P. (2024). Emotion AWARE: An artificial intelligence framework for adaptable and explainable emotion detection. Journal of Big Data, 11(53), 1–28. https: //doi.org /10.1186/s40537-024-00953-2

[12]. Ahmed, N., Farooq, U., & Shah, K. (2023). A systematic survey on multimodal emotion recognition: Acquisition methods, classification algorithms and fusion techniques. Array, 17, 100284. https://doi.org/10.1016/j.array.2022.100284

[13]. Rathee, G., Kumar, R., & Iqbal, R. (2023). A secure emotion-aware intelligent system for the Internet of Things. Future Generation Computer Systems, 148, 237–249. https: //doi.org /10.1016/ j.future. 2023.03.021

[14]. Bianco, S., Colombo, V., & Napoletano, P. (2021). A smart mirror for emotion monitoring in home environments. Sensors, 21(22), 7637. https: //doi.org /10.3390 /s21227637

[15]. Meng, W., Cai, Y., Yang, L. T., & Chiu, W. Y. (2021). Hybrid emotion-aware monitoring system based on brain waves for the Internet of Medical Things. IEEE Internet of Things Journal, 8(13), 11058–11069. https: //doi.org /10.1109 /JIOT.2020.3008983

[16]. Zhu, C. (2023). Research on emotion recognition-based smart assistant system. Journal of System and Management Sciences, 13(5), 210–225. https: //www. aasmr.org/jsms/Vol13/No.5/Vol.13%20No.5.15.pdf

[17]. Rajuroy, A., Johnson Mary, B., Cai, M., & Gu, D. (2025). Energy-efficient home automation via emotion-aware predictive modeling. International Journal of Smart Systems, 10(1), 1–14. https ://www. researchgate.net/publication/387820645

[18]. Shevhuk, I. (2025). Emotion-based voice control for IoT: Enhancing smart home automation. CEUR Workshop Proceedings, 3974, 44–52. https://ceur-ws.org/Vol-3974/short04.pdf