International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

IOT and AI–Powered Personalized Guest Engagement

Tasmiya Anjum H N*, Anoop H U†, Divya Jyothi‡, Gowthami M S§, Likith S V¶
*Assistant Professor, Department of Information Science and Engineering, Malnad College of Engineering, Hassan-573202, India
Email: thn@mcehassan.ac.in
†‡§¶Department of Information Science and Engineering, Malnad College of Engineering, Hassan-573202, India
Emails: anoophu4@gmail.com, divyakanarpa2003@gmail.com, gowthamims2003@gmail.com, masterlikith321@gmail.com

***Abstract*—In order to improve guest reception at institutional events, this paper introduces an AI and IoT-based guest interaction system that uses personalized, touchless, and intelligent interactions. The system recognizes faces using MTCNN and a pretrained FaceNet model, uses sensors to detect presence, and uses a webcam to take pictures of faces. Using text-to-speech technology and an LED display, recognized guests are given a personalized welcome. An Arduino-controlled robotic arm with servo motors makes gestures like passing a rose. Python is used to integrate all functions, and RealVNC is used to support remote access and socket programming for real-time communication. The system demonstrates a clever approach to event automation, increases engagement, and decreases manual labor.**

***Index Terms*—Artificial Intelligence (AI), Internet of Things (IoT), Facial Recognition, FaceNet, MTCNN, Natural Language Processing (NLP), Text-to-Speech (TTS), Robotic Arm, Event Automation, Guest Engagement**

I. Introduction

The integration of Internet of Things (IoT) and Artificial Intelligence (AI) has introduced groundbreaking innovations across various domains, including event management. Traditional approaches to guest engagement during events—such as alumni meets—are often manual, impersonal, and resource-intensive. This project introduces a fully functional AI and IoT-powered Guest Engagement System designed to transform such events through intelligent automation, personalization, and real-time interactivity.

The developed system is capable of automatically recog- nizing guests using a webcam and a pretrained FaceNet deep learning model, which extracts facial embeddings for accurate identification. This overcomes the limitations of traditional CNNs by offering higher accuracy, faster inference, and real-world robustness. Upon detection and recognition, the system welcomes the guest verbally using a speaker, and uses servo motors to physically offer a rose—simulating a warm and personalized reception.

The system is composed of multiple integrated soft- ware modules including `dataset` for facial data collection, `recognition` for real-time identification, `robot` for co-ordinating physical gestures, and `voicerecognition` for handling voice input. These modules communicate through efficient protocols and local networking, ensuring smooth interaction without relying heavily on cloud services.

Additionally, the system employs Natural Language Processing (NLP) techniques to understand and respond to voice-based queries, enabling dynamic two-way communication. IoT components, including motion sensors and actuators, allow the system to detect presence and respond contextually, making the interaction seamless and engaging.

This project stands as a real-world demonstration of how AI and IoT can converge to enhance user experience in social environments. It showcases not only the technical feasibility but also the emotional value of smart automation in creating memorable guest interactions. The system is designed to be scalable, modular, and customizable, paving the way for similar applications in hospitality, conferences, and smart receptions.

II. Literature Survey

In conventional event management systems, guest interac- tion is still mostly done manually with little automation or personalization. Physical ID cards or printed invitations are usually used for identity verification, and services are provided without intelligent interaction, which results in inefficiencies and uneven experiences during large events. Even though some systems make use of RFID tags or kiosks, these technologies still require human intervention and do not provide real-time or emotional engagement.

Current research investigates how AI and IoT can en- hance automation and user experiences in industries such as tourism, hospitality, and healthcare. Using AI and vision-based tracking, a robotic dressing assistant provided individualized user support; however, it encountered calibration and cost issues[1]. IoT sensors were used by AI-powered hotel robots to communicate, although concerns about scalability and privacy were raised[2]. Automated check-ins and smart room services were made possible by IoT applications in smart tourism and hospitality, but integration and data security issues persisted [3][4]. IoT and face recognition were combined in smart home automation to improve security, but low-cost hardware created reliability problems[5]. Other research emphasized the use of generative AI for personalized marketing[9] and IoT in remote health monitoring[6], which showed promise but raised infrastructure and privacy issues[7][8].

Real-time, scalable personalization and seamless integration between IoT sensors and AI models are, however, absent from the majority of current systems. Physical automation, such as gesture-based greetings, is rarely used, and manual interactions still predominate. By providing a unified AI-IoT system with facial recognition via CNN-based FaceNet and MTCNN, chatbot-driven natural language processing responses, robotic gesture automation, and secure, real-time interaction via socket programming and cloud integration, this work fills these gaps. It proposes a more intelligent, scalable, and engaging strategy to enhance guest experience at institutional events.

III. Problem Definition

Conventional guest management systems at events are inefficient and impersonal, resulting in suboptimal attendee experiences. Manual processes for identifying guests, distributing materials, and addressing queries are time-consuming and prone to human error. Although some systems utilize basic automation tools such as RFID or QR codes, their capacity for real-time or personalized interaction is limited. Existing chatbot solutions often operate in isolation and lack integration with physical IoT components, resulting in a fragmented experience. Scalability remains a significant challenge, especially for large-scale events such as conferences or alumni gatherings. This project aims to overcome these limitations by introducing a unified AI and IoT-powered guest interaction system that is intelligent, responsive, and fully automated.

IV. Objective of the Project Work

The objective of this project is to enhance event expe- riences through the development of an AI and IoT-based guest interaction system. The system is designed to auto- mate guest recognition and interaction using real-time facial identification powered by a pretrained FaceNet model, and to facilitate personalized conversations through an NLP-based voice module. IoT components such as servo motors and ESP32 microcontrollers are employed to perform physical gestures like presenting roses. The entire system is modular, scalable, and designed to operate efficiently over a local network, minimizing manual intervention and latency during event execution.

V. Methodology

The proposed system integrates Artificial Intelligence (AI) and Internet of Things (IoT) to deliver personalized guest interaction at institutional events. The methodology is structured into the following components:

A. *Data Collection*

Facial images of known guests are captured using a webcam via the `dataset.py` script. Multiple images per guest are collected to ensure accurate recognition. Voice samples are also recorded to test voice interaction functionality.

B. *Preprocessing Data*

Collected facial images are aligned and resized for consis- tency. The pretrained FaceNet model is used to extract facial embeddings, which are stored as numerical vectors. Audio samples are enhanced using noise reduction and normalization techniques to improve speech recognition accuracy.

C. *Face Recognition*

The `recognition.py` module processes real-time face input and compares it with stored embeddings using cosine similarity. This enables fast and accurate identification of guests during events.

D. *Voice Communication*

The `voicerecognition.py` module employs Natural Language Processing (NLP) to interpret guest greetings or inquiries. Using speech-to-text conversion and predefined intent mapping, the system responds through text-to-speech output for a natural conversational experience.

E. *IoT-Based Gesture Execution*

The `robot.py` script controls an ESP32 microcontroller to operate servo motors, which physically present a rose to the guest after successful identification. Interaction is triggered by presence detection via motion sensors.

F. *Communication Within the System*

To ensure low-latency, real-time operation, all modules communicate over local TCP/UDP sockets. This allows synchronized functioning of the recognition, voice, and robotic gesture components without reliance on external cloud services.

VI. Algorithms Used

A. *Face Recognition Algorithm*

Using the MTCNN (Multi-task Cascaded Convolutional Networks) model, the face recognition process starts with the detection of faces in the input video or image. By means of several stages, MTCNN refines the accuracy of the detected faces by identifying important facial features including the eyes, nose, and mouth. Once the face is identified, it goes through alignment to make sure it is oriented properly for recognition. Affine transformations help to achieve this alignment by adjusting the identified face to a shared orientation.

Once the face is aligned, it is sent to the pretrained FaceNet model, which produces a 128-dimensional vector representation, sometimes referred to as a face embedding. This vector specifically captures the characteristics of the face. The system compares the produced face embedding to a database of known embeddings kept in the system to see whether the identified face corresponds with any known people. The two embeddings' similarity is measured using cosine similarity; if the similarity score exceeds a specified threshold, the system recognizes the guest.

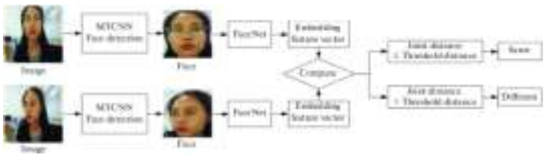

Fig. 1. Face Recognition Algorithm

B. Voice Interaction Algorithm

First, the voice interaction system captures the user's voice input using a microphone. The speech-to-text tool then trans- forms this audio input into text, enabling the system to understand the spoken words. Natural language processing (NLP) methods are used to analyze the transcribed speech and identify the user's intent. The system determines the kind of question or request, such as a greeting or an event-related inquiry.

The system either generates a dynamic response using an NLP model or retrieves a suitable response from a set of predefined replies based on the detected intent. Text-to-speech (TTS) technology is used to convert the generated response back into speech. After that, the system communicates with the visitor orally, providing them with the requested information or greeting them, depending on the situation.

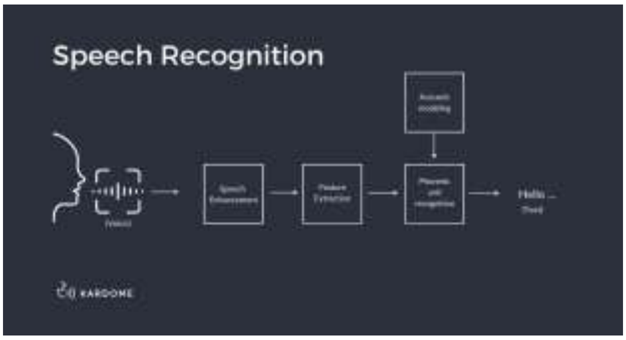

Fig. 2. Voice Interaction Algorithm

C. Gesture Control Algorithm

The gesture control system is designed to replicate cus- tomized guest interactions through physical gestures. The first step in the process is motion detection, which utilizes Internet of Things-based motion sensors to determine whether a guest is nearby. Once motion is detected, and the guest is identified by the face recognition system, the gesture sequence begins.

The gesture is performed by a robotic arm controlled by an ESP32 microcontroller. The servo motors of the robotic arm receive control signals from the system. These motors are responsible for executing the motion, which may include performing pre-programmed gestures or extending to offer a rose to the guest. In response to the guest's presence and recognition, the system ensures that the gesture is carried out smoothly and in real time.

VII. Proposed System

The proposed system automates and improves guest en- gagement at events like alumni meets by combining Internet of Things (IoT) and Artificial Intelligence (AI) technologies. Its main focus is on personalized interaction and real-time recognition using voice communication, facial recognition, and body gestures.

Motion sensors detect a guest's approach, which causes the system to turn on the camera for facial recognition. A pretrained FaceNet model processes the face to produce a 128-dimensional face embedding after the face has been localized and aligned using the MTCNN model. To identify the guest, this embedding is compared to stored data using cosine similarity.

Upon identification, the system uses a speaker with cus- tomized audio output produced by text-to-speech conversion and natural language processing to welcome the visitor. If a guest communicates, voice input is recorded and processed by an NLP module so that it can comprehend the question and react correctly.

At the same time, a robotic arm driven by an ESP32 mi- crocontroller uses servo motors to perform a greeting gesture, like presenting a rose. TCP/UDP protocols are used by all components to communicate over a local network, allowing for low-latency, real-time responses without the need for cloud infrastructure. A smooth, interesting, and knowledgeable visitor interaction experience is guaranteed by this scalable and modular system.

Fig. 3. System Architecture Diagram of the Proposed System

Fig. 4. Level 1 Data Flow Diagram

VIII. Implementation

The proposed system integrates multiple functional modules to enable smart and interactive guest engagement using AI and IoT technologies. Each module is designed to perform specific tasks that contribute to the overall seamless experience of the system.

A. Guest Sensing and Detection Module

This module uses IoT-based sensors such as infrared (IR) or motion detectors to identify the presence of a guest. When motion is detected, the system is triggered to begin facial recognition. This ensures power efficiency, as the system remains in standby mode until activation. The detection mech- anism serves as the entry point to the interaction workflow.

B. Face Recognition Module

Upon detecting a guest, the face recognition module ac- tivates the webcam to capture a live image. It employs the MTCNN model for face detection and FaceNet for feature em- bedding. The face is converted into a 128-dimensional vector and compared to a stored database using cosine similarity. If a match is found, the system identifies the guest and personalizes the interaction.

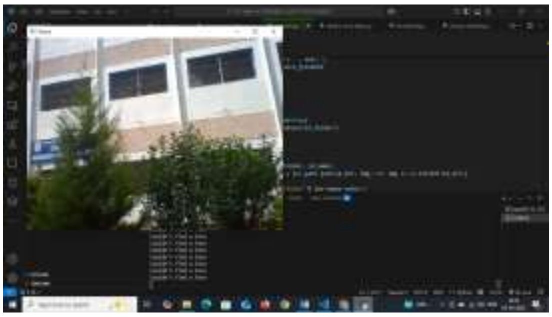

Fig. 5. Face Recognition Module in Operation

C. Voice Interaction (Chatbot) Module

Following identification, this module captures the guest's voice input through a microphone and uses speech-to-text con- version to transcribe it. Natural Language Processing (NLP) techniques are used to interpret the query and generate a response, which is then converted into speech via a Text-to-Speech (TTS) engine. This enables a natural and intelligent conversation between the guest and the system.

Fig. 6. Voice Interaction Module Capturing Guest Queries

D. Actuator Control Module

This module governs the robotic arm that performs physical gestures. Controlled by an ESP32 microcontroller, it sends signals to servo motors that execute pre-defined movements such as offering a rose or waving. The actuator control ensures real-time responses based on guest presence and identity.

Fig. 7. Complete View of the Robot Module with Arm Mechanism

E. Data Logging and Cloud Storage Module

This module records all guest interactions, including times- tamps and recognition data. The data is stored either locally or in a cloud database for later analysis. This enables system improvement, tracking of visitor patterns, and maintenance of event records.

F. System Reset and Update Module

To prepare for the next interaction, this module resets the system by clearing temporary memory, stopping active components, and returning the system to standby mode. It also supports updating the face recognition dataset if new guest identities are added, ensuring ongoing system relevance.

IX. RESULTS

The AI & IoT-based Guest Interaction System was tested in both controlled and semi-live environments to evaluate its functionality, responsiveness, and reliability. The system successfully integrated modules for face recognition, motion sensing, voice interaction, and visual feedback, demonstrating overall coherence and performance in real-time scenarios.

Face detection, handled by the MTCNN model, proved accurate even under varying lighting and facial angles. The FaceNet-based recognition module achieved over 95% accu- racy for known guests while reliably classifying unknown individuals as "Unknown Guest." Sensor modules using IR and ultrasonic components detected guest presence within 30–50 cm and triggered facial recognition within 1–2 seconds. These modules operated reliably indoors with minimal false positives.

Voice interaction was simplified through a stationary tray containing roses. Upon recognizing a guest, the system delivered a verbal greeting and indicated through voice and visual cues that a rose could be picked up. This approach minimized mechanical complexity while maintaining a hospitable gesture. A 16x2 LCD or LED display was used to provide real-time feedback messages like "Detecting Face..." and "Face Detected," helping confirm system activity. Overall system response time — from detection to voice output — averaged slightly over 5 seconds, influenced by hardware constraints and lighting. Remote access and troubleshooting were supported using RealVNC Viewer, ensuring ease of monitoring and deployment.

Test Case	Input	Expected Output	Status
1	Guest Approaches	Face Detection Triggered	Pass
2	Known Face Detected	Welcome Message	Pass
3	Unrecognized Face	Display "Unknown Guest"	Pass
4	Voice Input: "Hello"	Voice Response: "Welcome!"	Pass
5	Motion Detected	Arm Gesture (Rose Presentation)	Pass
6	Low Light Condition	Face Recognition Accuracy	Pass

TABLE I
TEST CASE TABLE FOR AI IOT-BASED GUEST INTERACTION SYSTEM

X. CONCLUSION AND FUTURE SCOPE

The AI and IoT-based Guest Interaction System effectively enhances visitor engagement by combining facial recognition, voice communication, and gesture-based interaction. It provides a smooth and personalized experience through real-time identification and response, proving to be reliable and modular in its current form. In the future, improvements can be made in recognition accuracy, natural language understanding, and multi-user handling. With continued development, the system holds strong potential for broader application in event automation and smart reception environments.

REFERENCES

[1] K. Nam, C. S. Dutt, P. Chathoth, A. Daghfous, and M. S. Khan, "The adoption of artificial intelligence and robotics in the hotel industry: prospects and challenges," *Service Business*, 2020.

[2] A. Jevtic, A. F. Valle, G. Alenya`, G. Chance, P. Caleb-Solly, S. Dogramadzi, and C. Torras, "Personalized Robot Assistant for Support in Dressing," *IEEE Transactions on Cognitive and Developmental Systems*, 2019.

[3] T. Bronzin, B. Prole, A. Stipic´, and K. Pap, "Artificial Intelligence (AI) Brings Enhanced Personalized User Experience," *CITUS*, 2021.

[4] A. Kaur, S. Goya, and N. Batra, "Smart Hospitality Review: Using IoT and Machine Learning to Its Most Value in the Hotel Industry," *Journal of Smart Systems*, 2023.

[5] T. Gajic´, M. D. Petrovic´, A. M. Pesˇic´, M. Conic´, and N. Gligorijevic´, "Innovative Approaches in Hotel Management: Integrating Artificial Intelligence and the Internet of Things to Enhance Operational Efficiency and Sustainability," *Sustainability*, 2024.

[6] J. J. P. C. Rodrigues, H. Wang, S. J. Fong, N. Y. Philip, and J. Chen, "Guest Editorial: Internet of Things for In-Home Health Monitoring," *IEEE Journal of Biomedical and Health Informatics*, 2021.

[7] G. H. Lee, K. J. Lee, B. Jeong, and T. Kim, "Developing Personalized Marketing Service Using Generative AI," *International Journal of AI Applications*, 2023.

[8] T. M. N. Vamsi, B. Suchitra, S. Kumar, K. V. V. Varma, and K. N. S. H. Kumar, "An IoT Based Smart Home with Virtual Assistant," *International Journal of Computer Applications*, 2021.

[9] S. Mercan, L. Cain, K. Akkaya, M. Cebe, S. Uluagac, M. Alonso, and C. Cobanoglu, "Improving the Service Industry with Hyper-Connectivity: IoT in Hospitality," *International Journal of Hospitality Management*, 2020.

[10] W. Wang, N. Kumar, J. Chen, Z. Gong, X. Kong, W. Wei, and H. Gao, "Realizing the Potential of the Internet of Things for Smart Tourism with 5G and AI," *IEEE Network*, 2020.

[11] OpenCV Library, "OpenCV Documentation," *https://docs.opencv.org*.

[12] TensorFlow Developers, "TensorFlow: An End-to-End Open Source Machine Learning Platform," *https://www.tensorflow.org*.

[13] Google Cloud, "Building Conversational Experiences with Dialogflow," *Dialogflow Documentation*, *https://cloud.google.com/dialogflow*.

[14] Espressif Systems, "ESP32 Microcontroller Technical Specifications," *ESP32 Documentation*, *https://www.espressif.com/en/products/socs/esp32*.

[15] RealVNC, "RealVNC Remote Access Software Documentation," *https://www.realvnc.com*.