Gesture-Based Appliance Control for Dumb using IOT and Machine Learning

Veda Shree Y. C, P. Prasanna, S. K. Shivashankar, H. P. Mohan Kumar

Department of Computer Science and Engineering, PES College of Engineering, Mandya, Karnataka, India

Abstract: This research paper details the architectural design and practical deployment of a gesture-responsive appliance automation framework that fuses Machine Learning (ML) algorithms with Internet of Things (IoT) infrastructure to promote seamless and assistive control of intelligent domestic systems. The solution integrates the VGG16 deep convolutional neural network to perform classification of dynamic hand gestures acquired in real time through an integrated webcam module, mapping each recognized gesture to its corresponding device control signal. These signals are wirelessly communicated to an ESP32-based microcontroller, which regulates the operation of end devices including illumination units, ventilation systems, alert mechanisms, water pumps, and electronic access locks via relay-based actuation. This vision-oriented, non-tactile interface overcomes the usability limitations of conventional human-machine input methods such as mechanical switches, handheld controllers, and voice-activated assistants—particularly enhancing accessibility for users with motor or verbal impairments. The deployed classification model achieves an accuracy rate exceeding 94% and demonstrates high-efficiency gesture tracking in real-time conditions. Designed with a modular software-hardware architecture, the system enables scalable integration, cost-efficient deployment, and adaptability across heterogeneous smart home configurations—ultimately fostering independence and universal access for all end users.

Keywords: Gesture Recognition, IoT, Machine Learning, VGG16, Appliance Control, Smart Home, Accessibility, Assistive Technology.

1. INTRODUCTION

The rapid development of smart home technologies has reshaped how users interact with everyday appliances, enabling control and automation through platforms like mobile apps, IR remotes, and voice-controlled systems. While these systems are effective for the general population, they often fall short for people with physical disabilities, limited mobility, or speech-related challenges. For individuals in these categories, typical interfaces are frequently impractical, inconsistent, or unintuitive. To resolve this barrier to accessibility, this research proposes a computer vision-based gesture interpretation system that integrates machine learning (ML) techniques and IoT infrastructure for seamless device control. By processing hand gestures recorded via webcam and classifying them with a fine-tuned VGG16 convolutional neural network, the system converts user intent into actionable instructions. These instructions are transmitted wirelessly to an ESP32 controller, which activates the appropriate appliance using relays. This method allows hands-free, speech-free, and contactless interaction, making it suitable for users who are unable to operate traditional controls or give voice inputs. It enhances the user's sense of autonomy and supports inclusivity, especially in assistive or caregiving settings. The system's design is modular, affordable, and supports real-time performance, making it an applicable and efficient solution for modern smart living environments. Introducing this approach contributes toward developing universally accessible smart technologies without sacrificing operational effectiveness. This paper presents the architecture, workflow, implementation, and evaluation results that demonstrate the practical viability of the proposed model.

2. RELATED WORK

Gesture recognition has been an active area of research in the fields of computer vision, human-computer interaction (HCI), and assistive technologies. A wide array of strategies have been formulated for capturing and interpreting gestures to manage electronic devices, including sensor-based wearable systems and vision-driven deep learning solutions. This section reviews the contributions of previous studies and highlights the limitations they present in the context of accessibility and real-time home automation.

Smruti Kshirsagar et al. (2020) introduced a glove-based system that allowed elderly users to control home appliances using embedded sensors in wearable gloves. While accurate in its operation, the dependency on wearable devices may result in discomfort over extended periods and limit usability for people with motor challenges.

Ayaka Hatori and Hiroyuki Kobayashi (2017) developed a gesture recognition interface that used traditional machine learning classifiers such as Support Vector Machines (SVM) and k-Nearest Neighbours (k-NN) with

handcrafted features. These approaches, while computationally lightweight, lacked the robustness and generalization capabilities of deep learning models.

Sulthana Manal et al. (2020) explored convolutional neural network (CNN)-based gesture classification integrated with IoT platforms to enable appliance control. Their work laid the foundation for using deep learning in this domain but was limited to predefined environments and lacked modularity for real-world applications.

Many earlier implementations rely on overly complex sensor configurations or do not achieve consistent real-time responsiveness. The system proposed in this paper addresses these challenges by offering a vision-based, contactless, and scalable gesture recognition framework integrated directly with an IoT-enabled microcontroller (ESP32), ensuring accessibility, affordability, and ease of deployment for smart home environments.

3. EXISTING SYSTEM

Most conventional systems depend on either physical touch or voice commands to function, which presents accessibility challenges for users with mobility or speech impairments, limiting their ability to operate such systems independently.

- Manual switches: Require physical access, not suitable for all users.
- Infrared remotes: Limited by range and line-of-sight.
- Voice assistants: Require clear speech and can be unreliable in noisy environments.

These systems present accessibility issues and do not support non-verbal or motion-restricted users. Moreover, they often lack customization and flexibility in adapting to individual needs.

4. PROPOSED SYSTEM

The design of the proposed system incorporates an ESP32 microcontroller for control and communication, a relay driver module to interface with electrical appliances, a graphical user interface (GUI) for user interaction, and a deep learning-based VGG16 Convolutional Neural Network (CNN) model for accurate gesture recognition.

Advantages:

- No need for wearable devices.

- Operates in real-time.

- Cost-effective and scalable.

System Components:

- Webcam for gesture input

- VGG16 model for classification

- ESP32 microcontroller for device control

- Relay circuits for switching appliances

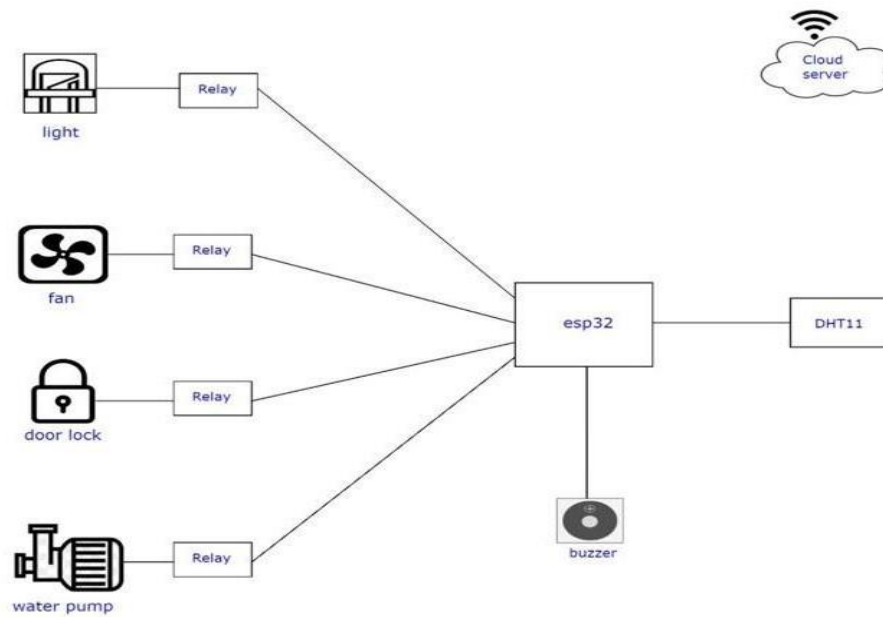- Python and OpenCV for image processing

Figure 1: System Architecture

5. METHODOLOGIES

Data Collection: Hand gesture images were gathered in different environmental conditions, including variations in background scenes and lighting intensities, to ensure the robustness and generalization of the gesture recognition model.

Preprocessing:
- Resize images to 224x224 pixels (input size for VGG16)
- Normalize pixel values
- Augment images for robustness

Model Training:
- Load pre-trained VGG16 model
- Replace the final layers with custom layers for classification
- Train on five gestures: light, fan, pump, buzzer, and lock

IoT Communication:
- Commands are transmitted via MQTT or HTTP
- ESP32 reads the command and triggers the respective GPIO pin
  Interface Design:
- Python GUI displays real-time video feed
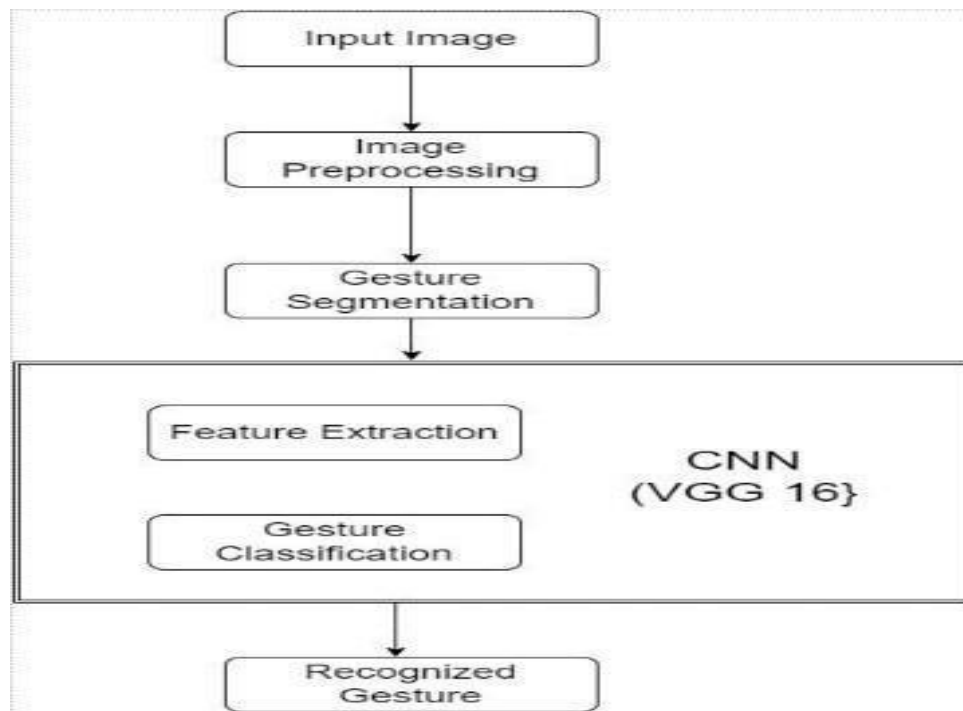- Detected gesture and action are displayed for feedback

Figure 2: Flowchart of the Proposed System

System Modules:
1. Gesture Detection – using OpenCV to extract the region of interest
2. Gesture Classification – using VGG16 for image classification
3. Command Mapping – mapping output class to IoT command
4. Transmission – sending data from PC to ESP32
5. Device Actuation – activating device based on received signal

6. WORKING OF SYSTEM

The system continuously captures video through a webcam. Each frame is processed to detect hand gestures. The extracted image is passed to the VGG16 model which predicts the gesture class. This classification is translated into an IoT command and sent via Wi-Fi to the ESP32. The ESP32, upon receiving the command, switches ON/OFF the corresponding appliance through a relay. Feedback is displayed to the user for confirmation.
Step 1: Real-Time Gesture Capture
Step 2: Image Preprocessing
Step 3: Gesture Recognition using VGG16
Step 4: Command Mapping and Transmission
Step 5: Appliance Actuation via ESP32
Step 6: User Feedback and Loop Continuity

7. RESULT AND DISCUSSION

Model Accuracy:
- VGG16 achieved an accuracy of 94.3% on the custom gesture dataset.
- Model showed good generalization across lighting variations and backgrounds.
  Hardware Performance:
- Commands were received by ESP32 within 0.5 seconds.
- Relay switching was accurate and reliable for all test cases.
  Testing Approach:
- Unit Testing: Verified ML classification output.
- Integration Testing: Verified command delivery from software to hardware.
- User Testing: Evaluated by users with restricted mobility.
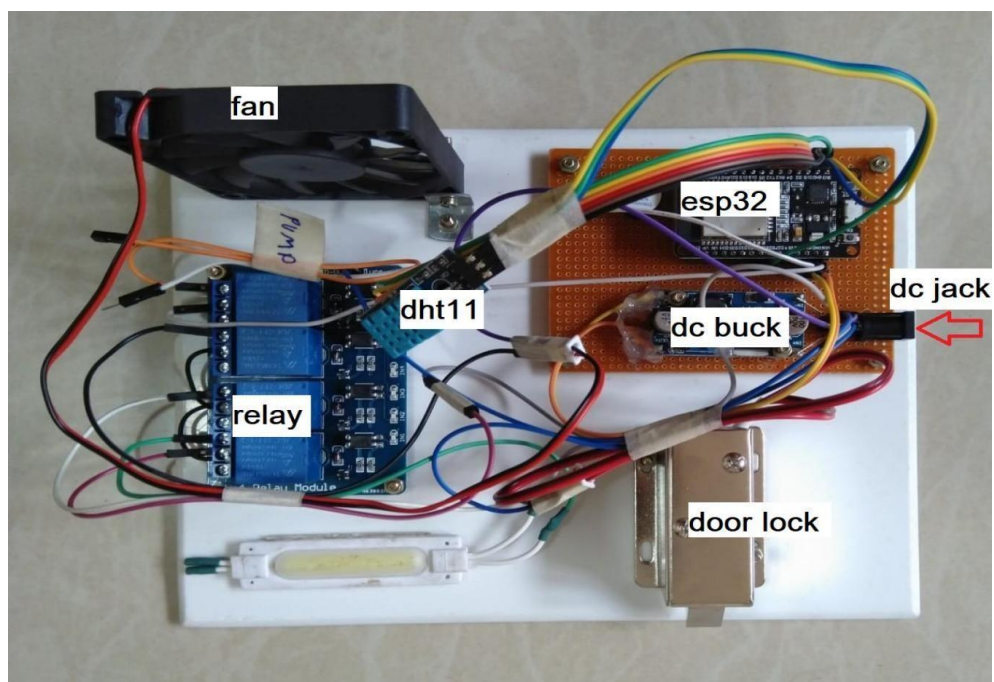
IJERTV14IS080012

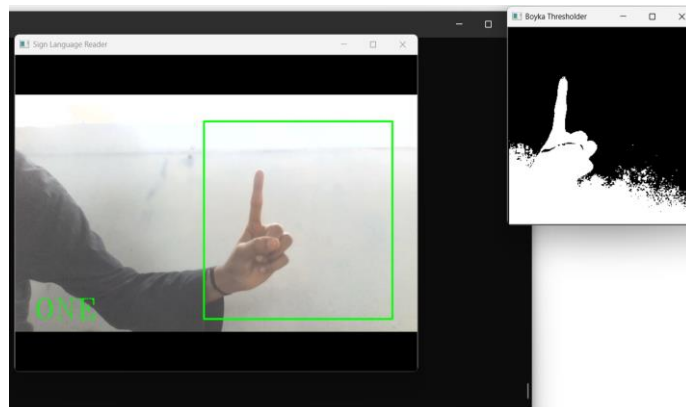Figure 3: Final Hardware Implementation

Observations:

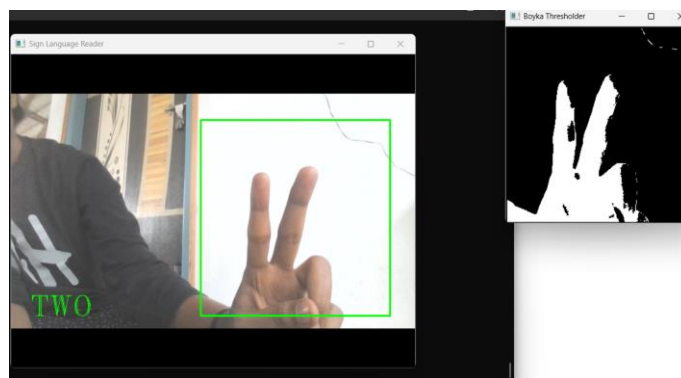Figure 4: Gesture 1 for Buzzer

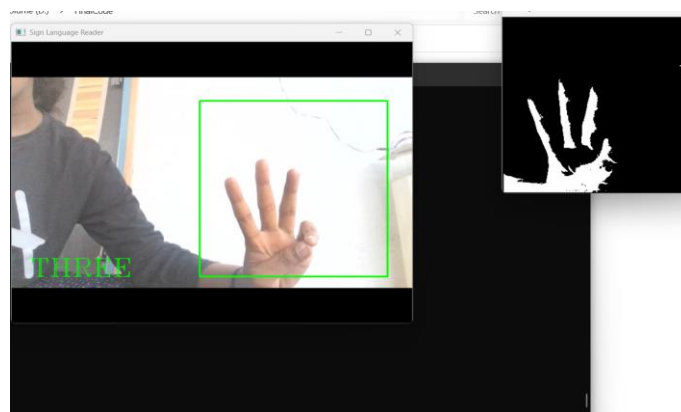Figure 5: Gesture 2 for DC Pump
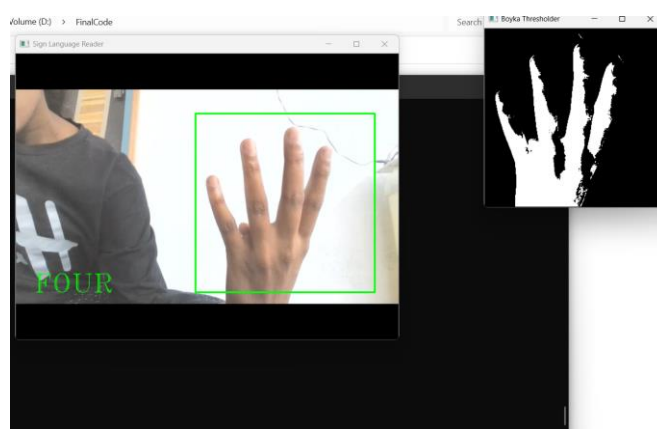
Figure 6: Gesture 3 for DC Fan
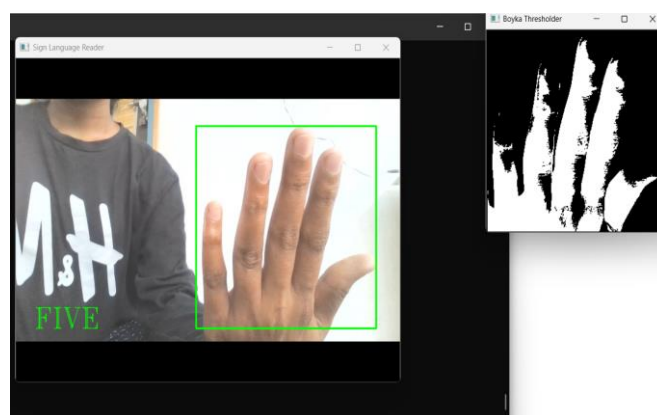

Figure 7: Gesture 4 for Light

Figure 8: Gesture 5 for Door Lock

- Users preferred the non-contact interface.
- Minimal training required to use the system.

8. CONCLUSION

This research introduces a novel and accessible solution for smart home automation that uses gesture-based control supported by machine learning and IoT. By combining the VGG16 convolutional neural network for gesture classification with the ESP32 microcontroller for appliance management, the system removes the dependency on physical switches or verbal inputs — making it particularly advantageous for individuals with physical limitations or speech challenges. The system achieved high recognition accuracy (above 94%) during real-time testing and effectively operated various appliances (light, fan, pump, buzzer, door lock) with negligible response delay. Its touch-free, vision-based interaction design ensures user convenience, while its modular system framework supports future customization and scalability. The validated performance of the system confirms its suitability for practical use in smart households, healthcare facilities, and environments designed for assistance. Enhancements in the future may include expanding the set of supported gestures and integrating facial recognition features.

9. REFERENCES

[1] S. Kshirsagar et al., "IoT Enabled Gesture-Controlled Home Automation," ICCMC, 2020
[2] T. H. Le et al., "IoT-based Hand Gestures Using Wearable Sensors," MAPR, 2019
[3] A. Hatori, H. Kobayashi, "IoT Device Control Using Gesture Recognition," SICE, 2017
[4] N. Jangamreddy, "Web-Based Gesture Recognition for IoT Control," COMSNETS, 2019
[5] S. Manal et al., "Catch Gesture: Recognition Technology with IoT," ICICCS, 2020