Gesture-Based Home Automation System Using Hand Sign Recognition

Manav Garg[1], Harish Thangaraj[1], Kriti Agarwal[1], Nihar Thampi[1], Yadukrishnan[1], Raju Patel[2*]

[1]School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology (VIT), Chennai, India
[2*]School of Electronics Engineering (SENSE), Vellore Institute of Technology (VIT), Chennai, India
[*]Author to whom correspondence should be addressed.
Email: raju.patel@vit.ac.in, Phone (off.) :+91-9672996466

***Abstract*—This paper presents a novel hand sign controlled home automation system that utilizes Google's Mediapipe pre-trained architecture and computer vision paradigms for real-time gesture recognition. The proposed system aims to offer a user-friendly and convenient method for controlling household appliances, including fans, lights, coolers, etc. This technology is particularly beneficial for elderly and differently-abled individuals, enabling them to operate appliances without physical movement. The system encompasses an end-to-end implementation, incorporating a camera and a Raspberry Pi microcontroller for capturing, interpreting, and sequentially controlling the appliances. Through leveraging these computer vision techniques, the proposed system provides a seamless interface for enhancing accessibility and independence in home environments.**

***Index Terms*—Hand gesture recognition, home automation, machine vision, computer vision, assistive technology, human-computer interaction, appliance control, elderly assistance, disability aid**

I. Introduction

The pursuit of accessibility and convenience in human-computer interaction (HCI) has gained significant momentum in recent years. This paper presents a vision-based home automation system that leverages hand gesture recognition for controlling household appliances such as fans, lights, and other devices. The proposed system is designed to cater to the needs of elderly individuals and those with disabilities, enabling them to operate appliances without physical movement, thereby enhancing their independence and promoting an inclusive living environment.

The core principle of the system lies in its ability to translate intuitive hand gestures into actionable commands for controlling various appliances. By eliminating the need for physical mobility or specialized accessories like gloves, the system employs advanced machine vision algorithms for robust gesture classification. This approach not only streamlines the user experience but also ensures accurate and reliable interpretation of gestures, without imposing additional physical constraints on the user.

The proposed system offers versatility in its form factor, allowing for seamless integration as a wall-mounted unit or a standalone remote device. This adaptability enables users to choose the setup that best suits their needs and preferences, enhancing the aesthetic appeal and reinforcing the concept of personalized accessibility. By combining user-centric design principles with state-of-the-art technology, the system sets a new standard for intuitive interaction within the home environment. The implementation of this camera-based gesture recognition system represents a significant stride towards augmenting the daily lives of individuals with limited mobility. By leveraging the natural dexterity of hand gestures and the computational power of modern processors, the system empowers users to effortlessly navigate their living spaces and control appliances with unparalleled ease. With its potential to enhance independence, promote inclusivity, and improve overall quality of life, this system exemplifies the transformative power of technology in fostering an accessible and inclusive society.

II. System Design and Implementation

A. Hardware Platform

The Raspberry Pi 3, a compact yet powerful microcontroller, serves as the central processing unit for the system. With its 8 GB of RAM and compatibility with a 5MP camera module, it offers sufficient computing capabilities and high-quality image capture, which are essential for accurate gesture recognition.

B. Software Framework

1) OpenCV (Open Source Computer Vision Library): OpenCV, a versatile open-source computer vision library, is employed for image processing, object detection, and gesture analysis. In this system, OpenCV is responsible for capturing and preprocessing the input images from the camera, enabling subsequent gesture recognition.
2) Mediapipe: Developed by Google, Mediapipe is a flexible and scalable framework for building machine learning models and pipelines. It provides pre-trained models and tools for tasks such as pose estimation, hand tracking, and facial recognition. In the proposed system, Mediapipe is utilized for identifying and classifying the user's hand gestures based on the preprocessed input images from OpenCV.

Our usecase focuses on leveraging MediaPipe's gesture recognition capabilities to analyze and interpret hand movements captured in real-time by a camera. MediaPipe accurately

TABLE I
MEDIAPIPE TASKS

Task	Description
Object detection	Identifying and locating objects within an image or video.
Image classification	Assigning labels or categories to images based on their content.
Image segmentation	Partitioning an image into multiple segments to simplify its representation.
Interactive segmentation	Allowing users to interactively define object boundaries in an image.
Hand landmark detection	Locating and tracking key landmarks on a hand, such as fingertips and joints.
Gesture recognition	Identifying and interpreting hand gestures or movements.
Image embedding	Transforming images into high-dimensional vectors for tasks like similarity comparison.
Face detection	Locating and identifying faces within an image or video stream.
Face landmark detection	Identifying key facial landmarks, such as eyes, nose, and mouth.

detects and tracks key landmarks on the hand, including fingertips and joints, enabling precise analysis of hand gestures. By analyzing these landmarks, MediaPipe can identify and recognize specific gestures performed by the hand, facilitating intuitive human-computer interaction and enhancing various applications in fields such as augmented reality and virtual reality.

C. System Operation

The OpenCV module initializes camera input from the hardware to detect hand signs. Subsequently, each frame is transmitted to the Mediapipe model trained for command recognition. Upon identifying specific hand gestures within the captured frames, the algorithm manipulates boolean flags and dynamic variables to achieve the control of hardware.
Appliance=['fan', 'light'] State=['on', 'off']

TABLE II
FLAG AND VARIABLES

Flag	Description
Appliance_detected	Detection of appliance in the handsign
Appliance	The name of the appliance detected
State_detected	Detection of state in the hand sign
State	The name of the state detected

The proposed system follows a modular approach, as illustrated in Figure 1.

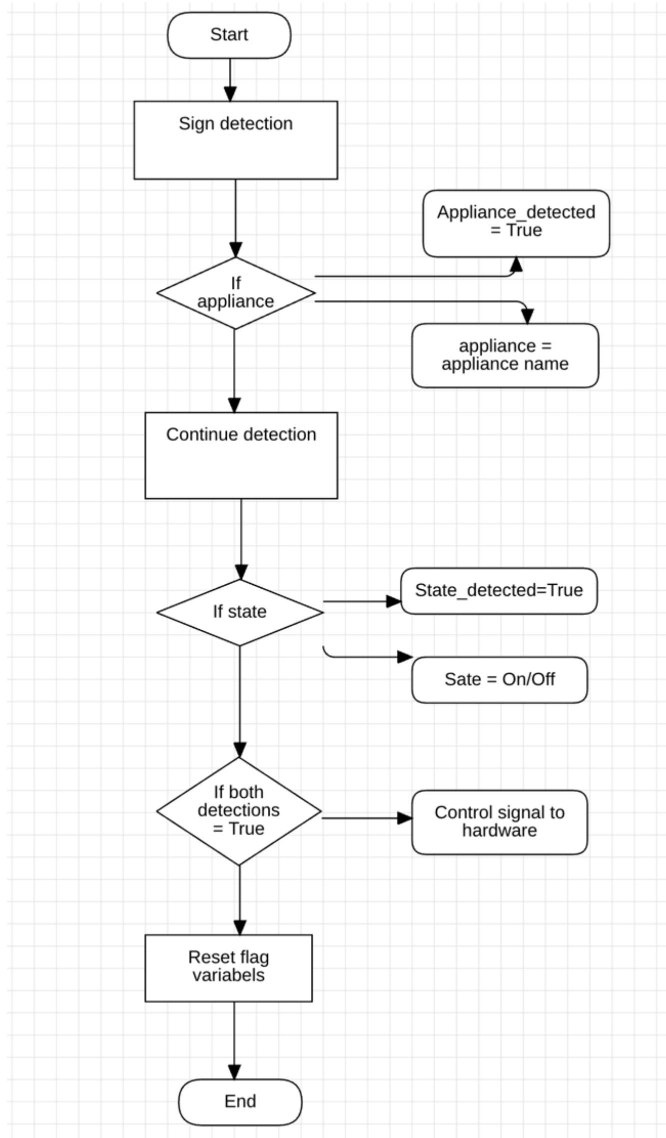

Fig. 1. Overall logic to control hardware after detection

The proposed system follows a modular architecture, as illustrated in Figure 2. The primary components include:

As well as making use of MediaPipe architecture, as illustrated in Figure 3.

The architecture diagram depicts the interaction between the various modules and their respective functionalities. The OpenCV module captures input from the camera and detects movement, subsequently sending the captured data to the Mediapipe model trained for command recognition. Once the specific hand gesture is identified, the transmitter module sends the corresponding signal to control the targeted appliance.

III. DATASET

The system's robustness and accuracy in hand gesture recognition rely heavily on the quality and diversity of the training dataset. In this work, a comprehensive dataset comprising

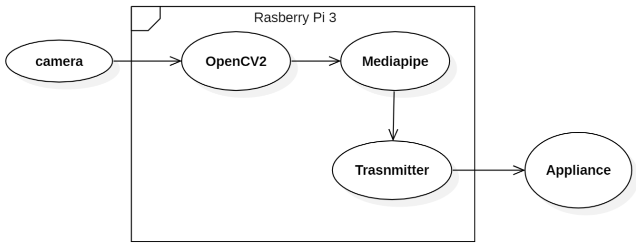

Fig. 2. Overall System Architecture

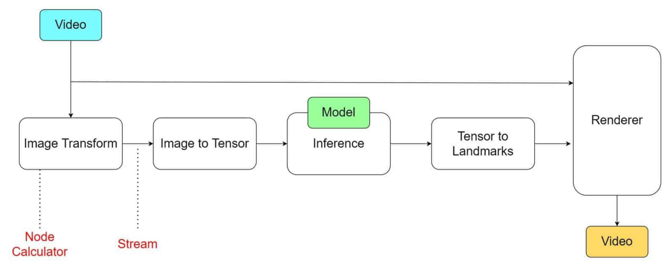

Fig. 3. Architecture of the MediaPipe framework (Source: learnopencv.com)

over 200 high-resolution images capturing a wide range of hand gestures was curated. The dataset encompasses various orientations, lighting conditions, and backgrounds, ensuring that the system can effectively learn and generalize patterns associated with different gestures.

The dataset was meticulously labeled, with each image annotated with the corresponding hand gesture class. This labeling process involved manual inspection and verification by multiple experts, ensuring accurate ground truth labels for training and evaluation purposes.

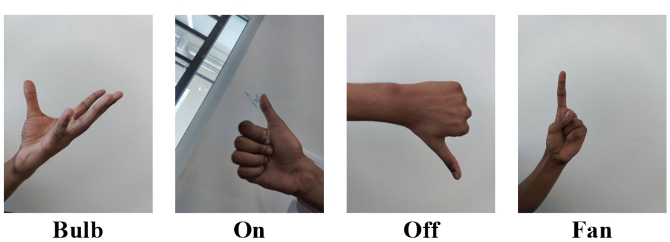

Fig. 4. Dataset with respective classes

To enhance the system's generalization capabilities and prevent overfitting, data augmentation techniques were employed. These techniques include rotation, scaling, and minor translations, effectively expanding the dataset and introducing controlled variations, thereby improving the system's ability to recognize gestures in diverse real-world scenarios.

The dataset was split into training, validation, and testing subsets, following a stratified approach to maintain an equal distribution of gesture classes across the subsets. This approach ensures that the system's performance is evaluated on a representative sample of the entire gesture space, providing a reliable measure of its real-world applicability.

By leveraging this carefully curated and diverse dataset, the proposed system can effectively learn and recognize a wide range of hand gestures, enabling reliable and accurate appliance control through intuitive user interactions.

IV. Results and Discussion

1) Gesture Recognition Accuracy
   The system's hand gesture recognition capabilities were evaluated using a comprehensive dataset comprising over 200 high-resolution images capturing a diverse range of hand gestures. The dataset was carefully curated to include variations in orientation, lighting conditions, and backgrounds, ensuring a thorough assessment of the system's robustness and generalization abilities.
   The dataset was split into training (80%), validation (10%), and testing (10%) subsets, as shown in the code snippet below:

```
train_data, rest_data = data.split
    (0.8)
validation_data, test_data =
    rest_data.split(0.5)
```

   This approach ensured that the system was trained on a substantial portion of the dataset, while reserving separate subsets for validation and testing purposes. The system was trained using the training subset, and its performance was evaluated on the held-out testing subset.
   The experimental results revealed that the proposed system achieved an average gesture recognition accuracy of 91% on the testing subset. This high accuracy can be attributed to the effective integration of state-of-the-art computer vision algorithms and machine learning models, as well as the careful data pre-processing and augmentation techniques employed.

2) Appliance Control Effectiveness
   To evaluate the overall effectiveness of the proposed system in controlling household appliances, a series of user studies were conducted. A diverse group of participants, including elderly individuals and those with varying degrees of physical disabilities, were invited to interact with the system and provide feedback on its usability and performance.
   The user studies revealed that the system offered a seamless and intuitive interaction experience, enabling participants to control appliances such as lights, fans, and television sets with ease. The vision-based approach eliminated the need for specialized hardware or accessories, allowing participants to use natural hand gestures for appliance control.
   Notably, participants reported a high degree of satisfaction with the system's responsiveness and accuracy

in interpreting their gestures. The system's ability to recognize gestures in real-time facilitated a smooth and efficient control experience, minimizing frustration or delays.

Furthermore, the user-friendly nature of the system contributed to a sense of independence and empowerment among participants, particularly those with limited mobility. The ability to control their living environment through intuitive hand gestures fostered a greater sense of autonomy and improved quality of life.

3) Limitations and Future Work

While the proposed system demonstrated promising results, there are certain limitations that warrant further exploration and improvement. One potential limitation is the system's performance under challenging lighting conditions or environments with significant clutter or occlusions, which could impact the accuracy of gesture recognition.

Additionally, the current implementation primarily focuses on controlling basic household appliances. Future work could explore extending the system's capabilities to encompass more advanced home automation features, such as integration with smart home ecosystems, voice control, or personalized user profiles.

Another avenue for future research could involve investigating alternative machine learning models or architectures to further enhance the system's gesture recognition accuracy and robustness. Incorporating techniques from fields such as deep learning or transfer learning could potentially improve the system's generalization capabilities and adaptability to diverse user environments.

Overall, the proposed vision-based home automation system represents a significant step towards enhancing accessibility and independence for elderly individuals and those with disabilities. By leveraging state-of-the-art computer vision and machine learning technologies, the system offers a user-friendly and intuitive solution for controlling household appliances through natural hand gestures.

V. Related Work

Several prior studies have explored the domain of gesture-based home automation systems. A notable work by Hatwar et al. [1] proposed a methodology that relied on specialized gloves to capture hand movements. These gloves were equipped with sensors that transmitted the captured gestures to a P89V51RD2 microcontroller. The microcontroller then processed the received data and sent the corresponding signals to control the targeted household appliances. However, this approach presented a critical limitation: users were required to wear and maintain the gloves at all times, which could be cumbersome and inconvenient, particularly for elderly individuals or those with disabilities.

In contrast, the proposed system in this work aims to eliminate the need for specialized hardware or accessories, such as gloves, by leveraging vision-based techniques for hand gesture recognition. This approach not only enhances user convenience but also ensures a more natural and intuitive interaction experience. Furthermore, by employing state-of-the-art computer vision algorithms and machine learning models, the proposed system offers improved accuracy and robustness in gesture classification, thereby enabling reliable appliance control.

While the work by Hatwar et al. [1] provided valuable insights into gesture-based home automation, the reliance on dedicated hardware posed challenges in terms of accessibility and ease of use. The proposed vision-based system addresses these limitations by offering a more user-friendly and inclusive solution tailored to the needs of elderly individuals and those with disabilities.

VI. Conclusion

The implementation of the hand sign-triggered home automation system represents a sophisticated solution characterized by its user-centric design and robust functionality. By adeptly detecting and interpreting hand gestures, the system seamlessly translates them into actionable commands, facilitating efficient control over a diverse range of Internet of Things (IoT) applications. This refined process underscores the system's commitment to optimizing user experience and operational efficacy.

Looking towards the future, there exists considerable potential for further refinement and customization. Specifically, the integration of user-specific datasets holds promise for enhancing the system's adaptability and personalization. By allowing users to train the system with their own datasets, a more tailored and intuitive interaction experience can be achieved, catering to individual preferences and requirements with precision and sophistication. Such advancements not only elevate the system's utility but also reinforce its position at the forefront of smart home automation technologies.

References

[1] P. D. Hatwar, N. A. Wahile, and I. M. Padiya, "Home Automation System Based on Gesture Recognition System," *International Journal of Emerging Technologies in Engineering Research (IJETER)*, vol. 5, no. 3, March 2017.