Visual Gesture Controlled Smart Home System

Mahitha M, Malavika A S, Nidhisha S Dharan, Shamnas S K
Department of Computer Science and Engineering, LBS Institute of Technology for Women, Poojappura, Thiruvananthapuram

Abstract: **The increasing integration of IoT devices in home automation systems has opened up new avenues for accessibility, particularly for individuals with physical, sensory, or intellectual disabilities. However, these advancements pose challenges for individuals who are mute, bedridden or rely on sign language, particularly when these devices are controlled by voice commands. This project proposes a gesture-based home automation system that utilizes hand gesture recognition technology, designed to enable effortless control of IoT devices through simple hand movements. Our system detects hand gestures in American Sign Language and translates them into commands to control various household appliances. To further enhance the accuracy and reliability of gesture detection, image enhancement using autoencoders and denoising techniques are integrated into the system, ensuring consistent performance even under varying lighting conditions and environments. This technology not only enhances accessibility for people with disabilities but also offers a more intuitive and contactless interface for all users via the gesture recognition using yolov8..Our system demonstrates high reliability and ease of use, with significant potential for further improvement through the integration of artificial intelligence and machine learning techniques.**

Index Terms—Gesture-recognition, Home automation system, American Sign Language(ASL), Autoencoders, Denoising,yolov8.

I. INTRODUCTION

In the expedited and fast-updating world, human living has become a transforming one with varying technologies. Among them, home automation is a unique and innovative idea that evolves as a life-changer in the lives of many people, particularly a visual gesture-controlled smart home system hard-of-hearing individuals communicate through sign-language to control appliances with hand gestures. Visual Gesture-Controlled Smart Home System uses computer vision and machine learning techniques to recognize ASL gestures and translate the identified gestures as commands for enabling the working of home appliances like lights, fans and electrical sockets. By translating hand movements into commands, the system enhances user-independence and convenience. This technology promotes a more inclusive and connected home environment. The project combines hardware components such as camera with software that processes visual data and recognizes gestures. It is tested for accuracy, responsiveness and usability in varying conditions which make sure the reliability and accessibility of the smart home system. In addition to that, the visual gesture based automation system addresses the challenges faced by the traditional systems which often rely on voice commands that makes the system less accessible to the people with motor disabilities or speech impairments.

II.LITERATURE REVIEW

Gesture recognition is a key generation in domestic automation, specifically for helping bedridden and disabled people. Static gestures, which contain studying a single frame, are more appropriate for such customers as compared to dynamic gestures that require monitoring a couple of frames. numerous studies papers discover specific gesture popularity techniques to manipulate domestic home equipment. A few structures, like [1] MobiGo, use a digital camera and CNN-based total models to locate hand gestures and send instructions to IoT gadgets. Others combine [2] multimodal user interfaces (MUI) that integrate voice and picture recognition for stepped forward accuracy. [3] Movement monitoring and gesture-based total interaction also play important roles in healthcare programs like physical remedy. Many answers make use of microcontrollers like [4] Arduino and Raspberry Pi, at the side of sensors including accelerometers and gyroscopes, to permit clever home management. Moreover, a few structures hire characteristic extraction strategies like [5] LDA and t-SNE to enhance gesture classification. To enhance accessibility and inclusivity, gesture-managed smart home structures had advanced the usage of numerous techniques. A few techniques rely on [6] smart gloves embedded with sensors to locate hand actions, changing them into textual content or speech, that's useful for deaf and mute people. different systems use switch getting to know and deep learning strategies like [7] CNNs and LSTMs to enhance recognition accuracy. [8] IoT integration performs an important function in permitting far off manipulation of devices, making home automation greater efficient and person-pleasant.actual-time gesture-to-textual content translation for signal languages like [9] ASL has additionally been explored, enhancing communication for human beings with disabilities. standard, those technologies offer low-value, powerful, and reachable answers to enhance the independence and pleasantness of life for disabled people.

III METHODOLOGY

The workflow of the gesture-controlled smart home system has been designed with the aim of real-time video processing for efficient, user-friendly controlling of household-appliances using hand gestures. A Raspberry pi module as the central processing unit.not only ensures a robust home automation system, but also make the system user-accessible. In the initial setup of hardware, the Raspberry Pi is connected to various components: a USB webcam captures real-time video, a relay module that maintains a stability between the Raspberry Pi and devices , and an LCD screen displays device status updates. The Raspberry Pi interfaces with a power supply, a motor driver for devices like fans, and GPIO wiring to manage device control signals. On the software side, essential libraries such as OpenCV for video capture, YOLOv8 for gesture detection, and GPIO control libraries are installed on the Raspberry Pi, creating a foundation for the gesture controlled automation system. This device empowers individuals with mobility impairments, presenting a user-friendly and available manner to control their surroundings. However, the device calls for a clear camera view and is sensitive to light situations that can affect gesture recognition. To enhance its abilities, the device may be improved to understand extra complicated hand gestures, thereby allowing the management of a greater variety of devices. Additionally, efforts may be made to enhance the device’s performance throughout specific light environments, making sure it really works correctly in any room or situation.

A. SYSTEM DESIGN

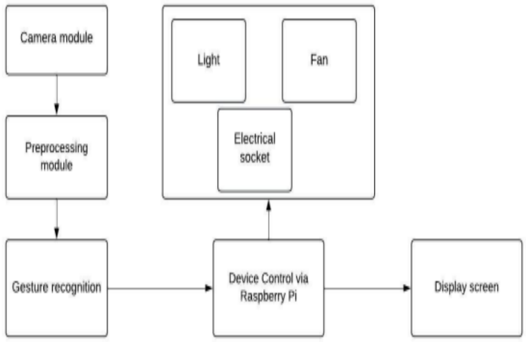


Fig 1: Overall system design with devices connected to Raspberry Pi

The Fig 1 shows the connectivity between different devices and the Raspberry Pi.The camera captures the video frames and then it undergoes preprocessing in the preprocessing module. After the preprocessing ,the frames is fed into the gesture recognition module for the gesture recognition using YOLOv8.The generated gesture labels in this module will converted into commands that will enable different appliances to work, where the device status will be displayed on a LCD screen.

B. CAMERA MODULE

The camera Module captures non-stop real-time video of the person's environment frame by frame, focusing on handy gestures. Each frame is processed without delay and sent for gesture recognition, making sure real-time tracking and on the spot reaction. In order to reduce the computational complexity, the camera only captures the even-numbered frames by frame skipping. This permits correct and efficient detection of hand gestures for seamless smart home control.

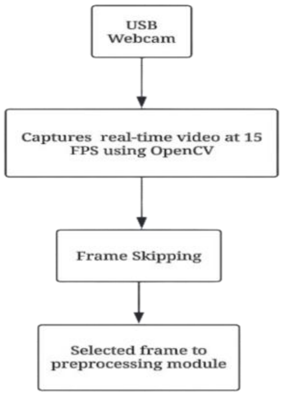


Fig 2: Camera Module

C. PREPROCESSING MODULE

The Preprocessing Module enhances the high-quality of captured pictures through adjusting brightness, enhancing contrast, and optimizing readability via autoencoders for higher gesture recognition. It reduces noise (denoising) to offer a cleanser input, making sure extra correct detection. The module additionally normalizes the picture, scales it to the specified dimensions, and applies filtering strategies to eliminate unwanted distortions. Moreover, it prepares the picture for the YOLOv8 version by means of refining edges, balancing color versions, and standardizing input formats, making sure unique and efficient gesture recognition.

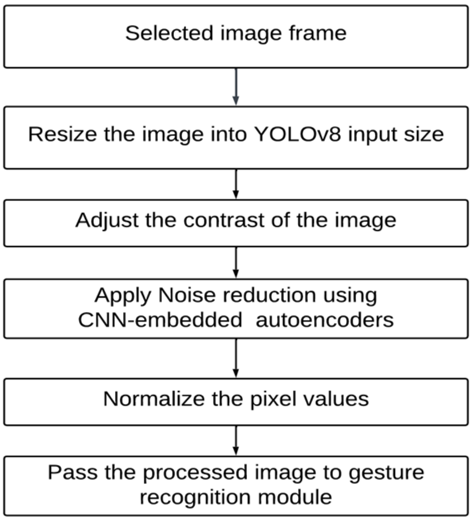


Fig 3: Preprocessing Module

D. GESTURE RECOGNITION

The Gesture recognition Module takes the preprocessed frame and feeds it into the YOLOv8 model, which is particularly trained to detect and classify hand gestures. YOLOv8 analyzes the frame in real time, figuring out the presence of unique gestures with excessive accuracy. The model is able to recognize gestures similar to unique sign language alphabets, including ASL (American sign Language). As soon as a gesture is detected, the device assigns a label to it, that's then used for similar processing. To enhance recognition accuracy, the module leverages superior feature extraction strategies, making sure dependable detection even in varying light situations. Moreover, it is able to differentiate among comparable gestures and handle more than one gesture input concurrently, making it strong for actual-global applications.

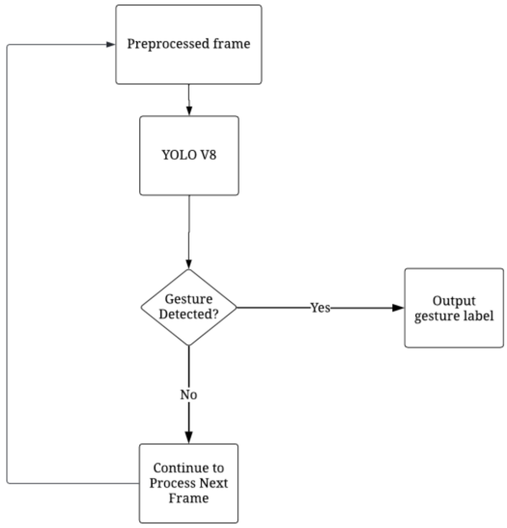


Fig 4: The Gesture recognition using YOLOv8.

E. DEVICE CONTROL VIA RASPBERRY PI

The device manager Module makes use of a Raspberry Pi to execute movements based totally on recognized hand gestures. As soon as the YOLOv8 model detects a gesture, it sends the corresponding label to the Raspberry Pi. The device then suits the label to a predefined action, such as turning lights on or off, adjusting fan speed, or activating other smart domestic gadgets. The Raspberry Pi communicates with home equipment the usage of GPIO

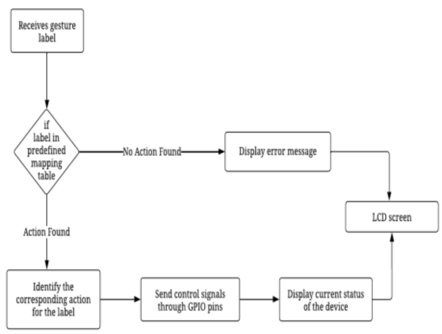
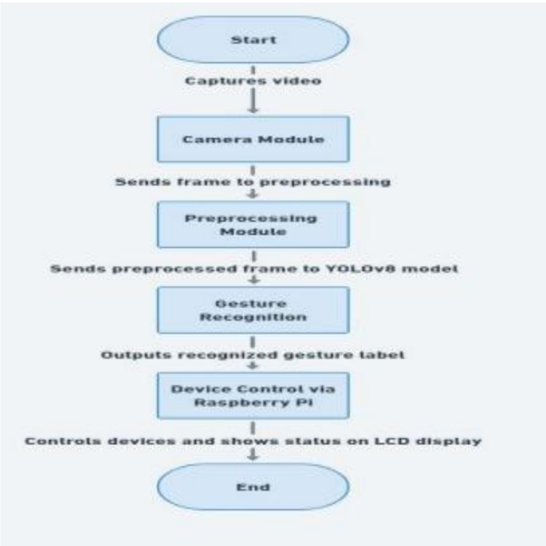


Fig 5: Device control using Raspberry Pi

(general-purpose input/output) pins, sending the necessary indicators to control them. Moreover, it is able to interface with relay modules to operate wireless and IoT-enabled devices. This permits for seamless, touch-free interaction with household home equipment, enhancing accessibility and comfort. Moreover, it presents real-time comments with the aid of updating an lcd display screen to reveal the modern status of the managed home equipment.

F. FLOWCHART

The system captures the real-time video as frames using a USB webcam.Then these frames are sent to the preprocessing module where they undergo several processes like image enhancement and denoising.The preprocessed frames would be directed to the gesture recognition module where the yolov8 model is used for training and recognition of captured hand gestures.Finally,it generates the labels for the corresponding gestures which goes to the raspberry pi module where the appliances are connected.In the Raspberry Pi module, the gesture labels are converted into control commands which, through the GPIO pins provide signal to the corresponding devices to work according to the commands. The LCD display which is connected to the Raspberry Pi will show the status of the devices connected.

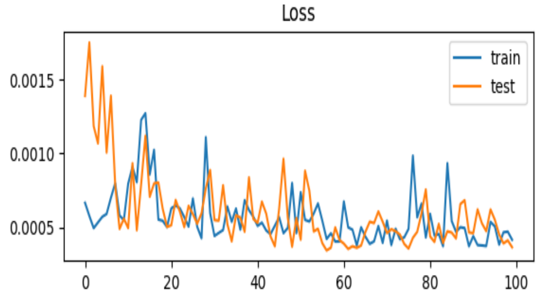


IV RESULTS

A. AUTOENCODER MODEL

The denoising autoencoder model was trained on noisy image data ,which uses a CNN based architecture of four convolution layers that consist with ReLU activations, batch normalization and skip connections inspired by ResNet, with mean squared error (MSE) as the loss function and Adam optimizer for quick stabilization. For evaluating the model performance, both the training and validation losses were monitored throughout the training process.

In the beginning, the model attained an accuracy of 56.73% with a loss of 0.016, but rapid progress was observed, with validation loss declining significantly from 0.03796 in the first epoch to 0.00030 at its peak performance. The model effectively learned to remove noise with an increase in validation accuracy peaking at 89.10%. Even though fluctuations in validation loss were observed during the later epochs, the model uniformly preserved fine image details. Though unstable at first, the model exhibited strong generalization accomplishing validation losses consistently below 0.001.The best-performing model was saved at 60 epoch with a validation loss of 0.00038 which demonstrated strong denoising ability.



The training and validation loss curves demonstrate a progressive drop over the epochs which indicates successful learning. The model achieved steady state after approximately 40 epochs, with both training and validation losses converging to lowest values. The relatively small gap between training and validation losses suggests that overfitting was effectively handled. Overall, the results confirm the capability of the model to denoise images effectively.

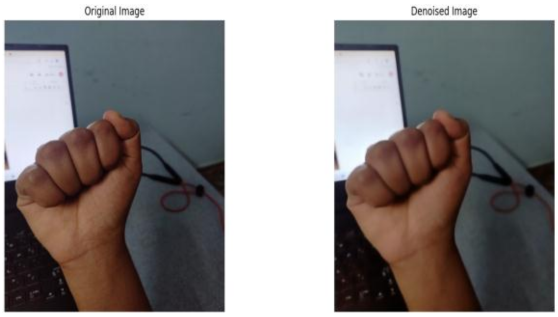


Fig 7: Prediction of Autoencoder

B. Gesture Recognition model

During training, the model underwent multiple epochs and its performance is monitored using key metrics such as training loss, validation loss, and mean average precision (mAP@50). The training loss curve indicates a continuous decline which reflects the ability of the model to learn effectively. Similarly, the validation loss curve also declined over time which denotes reduced overfitting. The mAP@50 score remained consistently high throughout training. The precision and recall values improved progressively.

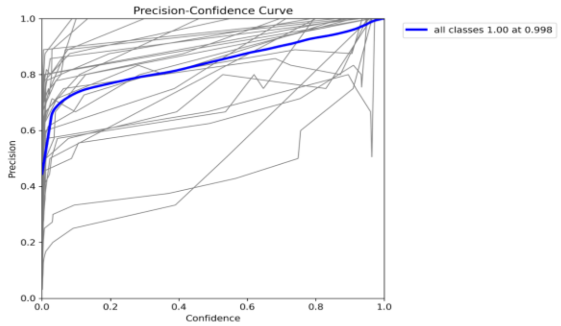


Fig 9: Precision-Recall Curve

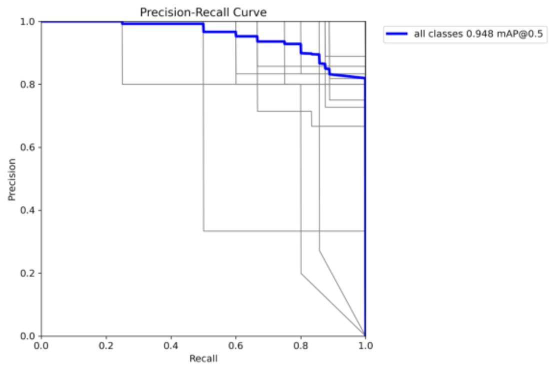


Fig 9: Precision-Recall Curve

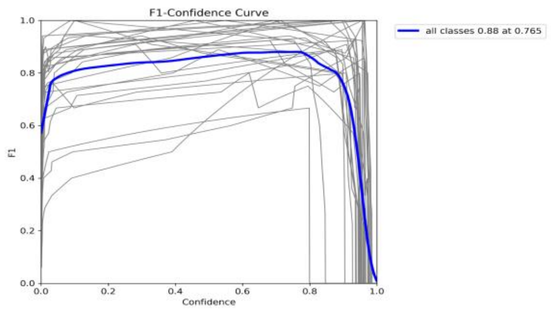


Fig 10: F1 Confidence Curve

The Precision-Confidence Curve demonstrates that it maintains high precision across various confidence levels, ensuring that most of the detected gestures are correctly classified. The Recall-Confidence Curve emphasizes the strong ability to identify actual positive detections by consistently achieving a recall score of above 0.9 at most thresholds. The Precision-Recall shows a slight drop in precision at higher recall levels. However, the overall performance remains strong, with an impressive mAP@50 score. The F1-Confidence Curve, with F1 scores close to 1.0,verifies that it maintains an optimal balance between precision and recall.###

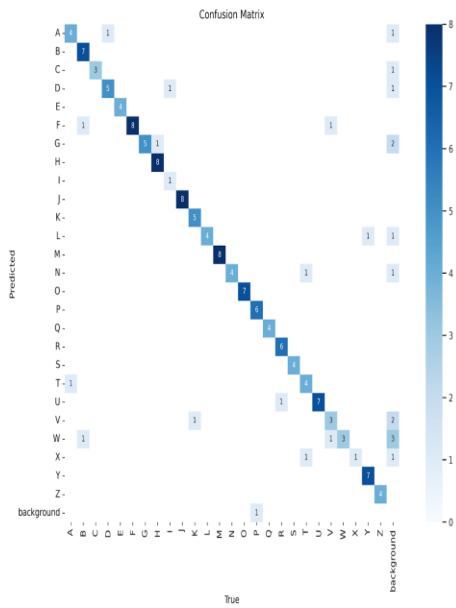


Fig 12: Confusion Matrix

The confusion matrix highlights both correct and incorrect classifications. Most predictions align well with their actual gesture classes, as seen in the **strong diagonal dominance. However, minimal misclassifications are observed among gestures with similar hand shapes.

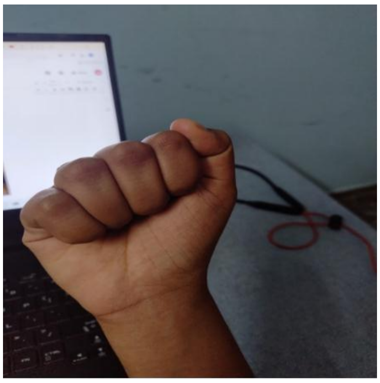


Fig. Input image of YOLO v8

```
Detected Labels:
Label: A, Confidence: 0.88
```

Fig 13: Output of YOLO v8

V CONCLUSION

In conclusion, the visual gesture controlled smart home system offers user-friendly , real-time control over home appliances through hand gestures.It pave the way into new scope for the people who face the challenge in communication and the individuals with disabilities in controlling household devices.By combining real-time video processing with advanced machine learning models like YOLOv8, the system can accurately detect and interpret hand gestures in ASL and , translating them into commands to control various devices. The optimized image preprocessing using autoencoders and gesture recognition techniques, maintains an appreciable accuracy and reliability of the system. Overall, this gesture-controlled automation system demonstrates a significant step forward in assistive technology, enhancing convenience and usability in daily life.

REFERENCES

[1] N. Jayaweera, B. Gamage, M. Samaraweera, S. Liyanage, S. Lokuliyana and T. Kuruppu,"Gesture Driven Smart Home Solution for Bedridden People," 2020 35th IEEE/ACMInternational Conference on Automated Software Engineering Workshops (ASEW),Melbourne, VIC, Australia, 2020.

[2] M. Alsaif and A. Albarrak, "Voice Command and Hand Gestures for SmartHome," 2020 First International Conference of Smart Systems and Emerging Technologies (SMARTTECH), Riyadh, Saudi Arabia, 2020.

[3] J. Xu, P. Mohan, F. Chen and A. Nürnberger, "A Real-time Hand Motion Detection System for Unsupervised Home Training," 2020 IEEE International Conference onSystems, Man, and Cybernetics (SMC), Toronto, ON, Canada, 2020

[4] K. Rathi, V. Sharma, S. Gupta, A. Bagwari, and G. S. Tomar, "Home Appliances using IoTand Machine Learning: The Smart Home," 2022 International Conferenceon Computational Intelligence and Communication Networks (CICN), Al-Khobar, Saudi Arabia, 2022,pp.329332.doi:10.1109/CICN56167.2022.10008294.

[5] Muhammad Muneeb, Hammad Rustam, and A. Jalal, "Automate Appliances via Gestures Recognition for Elderly Living Assistance," Feb. 2023, doi:https://doi.org/10.1109/icacs55311.2023.10089778

[6] S. Kshirsagar, S. Sachdev, N. Singh, A. Tiwari, and S. Sahu, "IoT Enabled Gesture Controlled Home Automation for Disabled and Elderly," Proc. 2020 Fourth Int. Conf.on Computing Methodologies and Communication (ICCMC), pp. 821-826, 2020.doi:10.1109/ICCMC48092.2020.ICCMC-000152

[7] S. Shetty, E. Hirani, A. Singh, and R. Koshy, "Gesture-to-Text: A Real-Time Indian Sign Language Translator with Pose Estimation and LSTMs," Procedia Computer Science, vol. 235, pp. 2684–2692, 2024

[8] Kurian, H. P, J. Regi, T. Y. Mahesh, and D. John, "Visual Gesture-Based HomeAutomation," 2023 3rd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS), Kanjirappally, India,2023, pp. 286-290, doi: 10.1109/ACCESS57397.2023.10200895.

[9] D. Tomar et al., "Gesture-Controlled Home Automation for the Differently Abled: Enhanced Accessibility and Independence," 2023 ICTACS, Tashkent, Uzbekistan, pp. 1560-1564. doi:10.1109/ICTACS59847.2023.10389929.

[10] R. A. Urmee, N. S. Prome , and T. Ahmed, "Hand Gesture-Based Home Automation System," TENCON 2022 - 2022 IEEE Region 10 Conference (TENCON), Nov. 2022