Gesture Controlled Home Automation for People with Disabilities

Michael Sichilongo
School of Engineering and Technology,
Department of Computer Science and Information Technology
Mulungushi University
Kabwe, Zambia
michealsichilongo@gmail.com

Brian Halubanza
School of Engineering and Technology,
Department of Computer Science and Information Technology
Mulungushi University
Kabwe, Zambia
bhalubanza@gmail.com

Abstract- This study examined a Gesture-Controlled home Automation system integrated with Artificial Intelligence (AI) to address accessibility challenges faced by individuals with disabilities and older adults. These groups often encounter difficulties operating standard electrical appliances such as lights, fans, and door locks without assistance. By relying on natural, intuitive gestures such as hand movements, head tilts, or other routine body motions, the system enabled users to control household appliances remotely, removing the need for physical contact with switches or reliance on wearable devices. During the research, a camera and AI algorithms were used to detect and interpret distinct gestures, which were then mapped to various appliance controls. The results showed that this contactless approach improved user convenience and reduced the need for caregiving support, enhancing independence both at home and in office settings. User feedback indicated that the system's straightforward, intuitive interface promoted confidence in everyday tasks, while also minimizing physical strain. Moreover, the findings suggested that such inclusive technologies could encourage broader social engagement by allowing people with mobility limitations to manage their environment more autonomously. Overall, the study demonstrated that adopting an AI-driven, gesture-based control system can significantly improve daily life for individuals with disabilities and older adults, further contributing to equitable and accessible living environments.

Keywords: Gesture Recognition, Artificial Intelligence, Accessibility, Assistive Technology, Independent Living

I. INTRODUCTION

Gestures can be defined as non-verbal bodily movements that convey messages or commands without the use of spoken language. These gestures—often involving the hands, head, or other parts of the body—are commonly categorized by their intended function, such as conversational, controlling, manipulative, or communicative [1]. With ongoing advancements in Machine Learning (ML) and Gesture Recognition Technology, gestures now serve as a feasible input method for numerous applications, including the operation of home or office appliances [2], [3].

In Zambia, many individuals with mobility limitations continue to rely on traditional practices in which family members or caregivers manually operate electrical appliances—such as turning lights on or off—on their behalf. While this approach may function on a basic level, it presents significant drawbacks, including limited accessibility, barriers to employment, and increased dependence on caregivers. Socio-economic challenges further hinder widespread adoption of current assistive technologies, which can be prohibitively costly or overly complex. Consequently, companies are often hesitant to hire persons with disabilities due to the lack of suitable, affordable solutions. Given these constraints, there is an evident need for a cost-effective, user-friendly, and adaptable system that removes barriers to accessibility. Specifically, an interfaceless mechanism capable of remotely manipulating electrical switches would grant individuals the autonomy to control devices from a distance without direct physical contact.

This project therefore focuses on harnessing Artificial Intelligence (AI) and the Internet of Things (IoT) [4] to address accessibility issues experienced by disabled and elderly individuals in Zambia. The primary objective is to create a gesture-controlled home automation system that detects, recognizes, and interprets user gestures in real time, thus enabling seamless interaction with physical hardware. The proposed framework integrates computer vision with a machine learning model trained on a diverse set of gestures, while a camera module functions as the primary input device.

By designing a platform able to accurately map recognized gestures to the control of various appliances, this research aims to foster greater inclusivity and autonomy for those with limited mobility. Beyond offering practical benefits for domestic and corporate settings, it aspires to enhance socio-economic participation among individuals with disabilities, ultimately demonstrating how AI-driven gesture recognition can help bridge persistent accessibility gaps. The anticipated results include a noticeable reduction in caregiver dependency and a marked improvement in users' overall quality of life, potentially transforming not only Zambian households but also other contexts where similar challenges persist.

Moreover, the outcomes of this project will inform further research, guiding refinements of gesture-controlled systems for broader assistive technology uses. This approach envisions collaboration among government agencies, academic institutions, and industry partners to foster an ecosystem of inclusive innovation. By addressing immediate accessibility needs while shaping equitable progress, the project provides a robust groundwork for transforming assistive technologies in Zambia and similar contexts.

III. PROPOSED SYSTEM

The proposed Gesture-Controlled Home Automation system offers a smart, cost-effective, user-friendly, and accessible solution for individuals with mobility limitations. By integrating machine learning and computer vision techniques, this approach allows users to control a variety of household appliances and environmental settings using intuitive human gestures, eliminating the need for direct physical interaction with switches or devices. The fundamental objective is to minimize dependence on caregivers and maximize autonomy for individuals who face challenges in performing routine tasks due to impaired mobility or related constraints.

At the heart of this system is a Convolutional Neural Network (CNN), functioning as the primary Machine Learning Algorithm. The CNN is carefully trained on a robust dataset of gesture signs, thereby enabling it to accurately recognize and classify diverse gestures [9]. Once a user performs a gesture within the camera's field of view, the algorithm processes the captured video data in real time, identifying and interpreting the command associated with that particular gesture.

In addition to the CNN, a dedicated Controller, typically implemented via a microprocessor or microcontroller, serves as the intermediary between the gesture-recognition software and the connected appliances. Upon receiving the classification results from the CNN, the controller transmits the corresponding commands to relevant devices, such as electric fans or light bulbs. This step ensures that the user's intended action—whether turning on a fan, dimming a light, or performing another function—occurs promptly and accurately.

The Camera plays a pivotal role in this setup by continuously capturing real-time video data, which is then fed into the CNN for processing. This arrangement allows the system to operate seamlessly, as the camera and controller work together to carry out gesture-based commands without manual intervention.

Figure 1: Proposed System

Figure 1 depicts the system architecture, highlighting a computer system linked to an electric fan and a light bulb. When the user executes a specific hand or body gesture, the camera detects it, sends the captured image stream to the computer for analysis, and prompts the CNN to determine the correct action. Subsequently, the controller signals the relevant appliance to execute the required function, completing the interaction cycle.

IV. METHODOLOGY

The system was developed with a clear focus on its target users, particularly elderly and disabled individuals, by prioritizing gestures that are easy for them to perform. To ensure accurate representation of individuals with mobility constraints, researchers collected data from local communities and care facilities. This process included identifying specific gestures that older adults and people with physical disabilities such as those without hands or legs could perform comfortably. By taking these factors into account, the system's design and dataset more closely reflect the needs of its primary user population.

A. Data Collection

Images of various gestures were captured using a camera. These gestures were selected to ensure ease of use by individuals with mobility challenges [10].

B. Data Processing

The collected images were annotated to prepare them for use in the machine learning model.

 ISBN: 978-9982-95-500-3

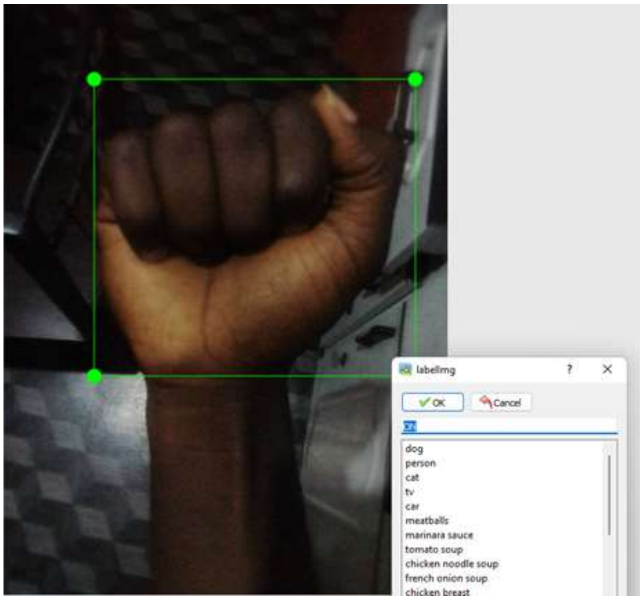

Figure 2: Image annotation

Figure 2 illustrates the image annotation process employed in this research. During this process, key regions corresponding to user gestures were manually labeled and classified in order to train the machine learning model. By accurately identifying and marking the boundaries of each gesture, the dataset became more informative, thereby improving the model's ability to recognize and interpret user actions.

C. Model Training

The model was developed using TensorFlow Lite's MobileNet, a pre-trained convolutional neural network (CNN) [11]. The dataset was divided into two parts, with 80% used for training and the remaining 20% reserved for testing.

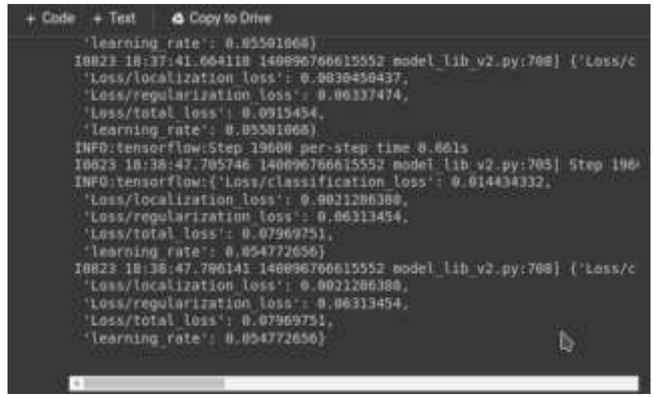

Figure 3: Training Metrics

The model's training performance was evaluated using metrics that included overall loss, classification loss, localization loss, and regularization loss as shown in figure 3. Low values in all these categories indicated that the model not only accurately recognized and localized objects but also avoided overfitting. After successful training, the model was integrated into a processing unit and a microcontroller, enabling intelligent control of home appliances through the recognized commands it processes.

V. RESULTS

The performance of the model during training was evaluated using key loss metrics, which provide insights into its effectiveness and accuracy. The overall loss was measured at 0.07968, indicating strong overall performance, as lower values suggest better training outcomes. The classification loss, which assesses the model's ability to correctly classify objects, demonstrated reliable performance, reflecting the model's capacity to differentiate between various gesture classes. Additionally, the localization loss was recorded at 0.00212, signifying high accuracy in identifying and locating objects within images. This low value highlights the model's precision in detecting relevant regions for gesture recognition. Finally, the regularization loss was 0.0631, which serves as a penalty term to prevent overfitting by discouraging overly complex model parameters. Collectively, these metrics reveal that the model trained successfully, as evidenced by the decreasing loss values over time, demonstrating its robustness and reliability for gesture-based object detection and classification tasks.

The trained model was integrated with processing unit and a microcontroller, which was programmed to control the connected home appliances based on command received from the processing unit.

 ISBN: 978-9982-95-500-3

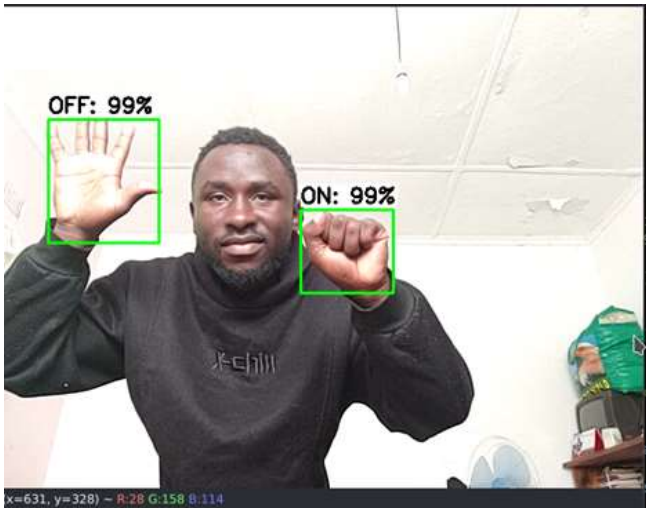

Figure 4: Image detection

Figure 4 demonstrates the results of the model's testing phase, where it successfully detected objects or states in an image and provided corresponding confidence percentages for each detection. The performance of the model is summarized as follows:

- OFF: 99% Confidence: The model identified an object or a state labeled as "OFF" with a high confidence level of 99%. This indicates the system's ability to recognize a device or appliance that is currently in the "off" state, showcasing its precision in understanding operational states.

- ON: 99% Confidence: Similarly, the model detected an object or a state labeled as "ON" with a 99% confidence level. This signifies the system’s capability to reliably recognize appliances or devices that are actively operating in the "on" state.

These results highlight the model's effectiveness in detecting and classifying states or objects with a high degree of accuracy. Such precise detection ensures that the system can confidently execute the appropriate commands for controlling home appliances, thereby enabling seamless and reliable interaction for the user. The integration of these detection capabilities into the broader system contributes significantly to its overall functionality and user-friendliness.

Performance Metrics

The model's performance was evaluated using standard object detection metrics, with results providing insights into its overall effectiveness and precision in identifying and classifying objects or states.

- mAP (mean Average Precision): The mAP metric serves as the primary measure of an object detection model's performance. It evaluates both recall (the proportion of correctly detected objects) and precision (the accuracy of those detections). The model achieved an overall mAP of 80.19%, reflecting a solid performance level for object detection tasks.

- Best Performance: Among the evaluated classes, the "ON" class demonstrated the highest performance, achieving a mAP of 89.28%. This indicates the model's strong ability to detect and classify objects or states associated with the "ON" label accurately.

- Worst Performance: The "PLAY" class showed the lowest performance, with a mAP of 55.00%, suggesting that the model faced challenges in accurately identifying or classifying objects in this category.

- IoU (Intersection over Union) Threshold: IoU measures the overlap between the predicted bounding box and the ground truth box, determining the accuracy of object localization. The model's performance was evaluated across IoU thresholds ranging from 0.5 to 0.95, demonstrating consistent and reliable detection accuracy across various levels of stringency.

These metrics highlight the model's capability to perform well in detecting and classifying objects under diverse conditions, while also revealing areas for improvement, particularly in the "PLAY" category. The solid mAP score and consistent IoU performance indicate that the model is well-suited for real-world applications where reliable object detection is essential.

VI. CONCLUSION

In conclusion, this study successfully developed and evaluated a gesture-controlled automation system designed to enhance accessibility, independence, and convenience for individuals with mobility challenges, including the elderly and those with disabilities. The system leverages advanced machine learning and computer vision technologies to recognize and interpret natural human gestures, enabling users to control various home appliances without the need for physical interaction. By providing a hands-free interface, this solution significantly reduces the barriers faced by individuals with limited mobility in managing their living environments.

The results of this study highlight the effectiveness and practicality of gesture recognition as a tool for promoting accessibility and inclusivity. The system’s ability to accurately detect and process gestures demonstrates its potential to empower individuals, fostering independence

ISBN: 978-9982-95-500-3

and reducing reliance on caregivers for routine tasks. This research also underscores the importance of integrating inclusive design principles into technological development, emphasizing the need to prioritize user-centric innovations that address the unique challenges faced by diverse populations.

As technology continues to evolve, the integration of gesture-controlled systems into everyday environments presents an opportunity to revolutionize accessibility solutions. From homes to workplaces, such systems have the potential to enhance the quality of life for people with mobility limitations by creating a more inclusive and independent living experience. This study serves as a foundation for future advancements in gesture recognition technologies, encouraging further research and innovation to broaden their impact and applicability. Ultimately, the findings advocate for a future where accessibility and independence are seamlessly embedded into the fabric of everyday life.

References

[1] S. Akyol, C. U. Bilgin, and H. Wirtz, "Gesture Control for Use in Automobiles," MVA, pp. 349-352, 2000.

[2] N. Altiok, S. N. Ünsal, "Design of a Smart Home System for Elders and Disabled People," in 2019 10th International Conference on Electronics, Computer Science, and Artificial Intelligence, pp. 1-4, 2019.

[3] H. S. Badi and S. H. Badi, "A Survey of Hand Gesture Recognition and Applications," International Journal of Computer Science and Engineering, vol. 6, no. 12, pp. 126-132, 2014.

[4] M. Bhuiyan and R. Picking, "Gesture-Controlled User Interfaces: What Have We Done and What’s Next," in Proceedings of the Fifth Collaborative Research Symposium on Security, E-Learning, Internet, and Networking (SEIN 2009), Darmstadt, Germany, pp. 26-27, Nov. 2009.

[5] El Naqa and M. Murphy, "What is Machine Learning?" in Machine Learning in Radiation Oncology, Springer International Publishing, pp. 3-11, 2015.

[6] S. Fels and G. Hinton, "Glove-Talk: A Neural Network Interface Between a Data-Glove and a Speech Synthesizer," IEEE Transactions on Neural Networks, vol. 4, no. 1, pp. 2-8, 1993.

[7] P. Halarnkar, S. H. Suryawanshi, S. H. Joshi, "Hand Gesture Recognition for Human-Computer Interaction," Berlin: Springer, 2012.

[8] Johnstone and E. de Dios, "The Role of Smart Home Technology in Reducing Social Isolation and Loneliness in Older Adults," Technology and Aging, vol. 1, no. 1, pp. 1-28, 2020.

[9] Marinelli et al., "Active Upper Limb Prostheses: A Review on Current State and Upcoming Breakthroughs," Progress in Biomedical Engineering, 2022.

[10]] Patil et al., "Sign Language Recognition Using Gesture Recognition and Natural Language Processing," in Data Management, Analytics, and Innovation: Proceedings of ICDMAI, pp. 197-210, 2021

[11] Zhai, S. H. M. a. S. B., 2000. The metropolis keyboard-an exploration of quantitative techniques for virtual keyboard design. In: Proceedings of the 13th annual ACM symposium on User interface software and technology. s.l.:s.n., pp. 119-128.