Grenze International Journal of Engineering and Technology, June Issue

Gesture Control in Healthcare System

Eswaraiah Rayachoti[1], Rajesh Duvvuru[2], Lakshmi Prasanna Meda[3] and Sudhir Tirumalasetty[4]
[1-3]SCOPE, VIT-AP University, Andhra Pradesh, India
Email: eswaraiah.rayachoti@vitap.ac.in, rajesh.duvvuru@vitap.ac.in, lakshmi.23phd7105@vitap.ac.in
[4]Department of AID, Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, India
Email: sudhir.t@vvit.net

***Abstract*— This research delves into the incorporation of gesture control technology within healthcare environments, emphasizing the augmentation of accessibility and convenience. The developed system empowers users, especially healthcare professionals, to utilize instinctive gestures for a range of tasks, including the regulation of environmental variables like air conditioning and lighting, as well as the management of doors and gates. Through the utilization of gesture recognition, the research seeks to optimize daily operations in healthcare facilities, offering a hands-free and streamlined approach to interact with the physical surroundings. The research scrutinizes the technical intricacies of implementing gesture control, assessing its efficacy in enhancing workflow efficiency and reducing physical contact. Consequently, the findings contribute to cultivating a healthcare environment that is not only more hygienic but also user-friendly.**

***Index Terms*— Gesture control, Healthcare**

I. INTRODUCTION

In recent times, the healthcare landscape has undergone a notable transformation through the integration of state-of-the-art technologies, ushering in innovative solutions to optimize efficiency and streamline daily operations. This research aims to actively contribute to this ongoing paradigm shift by delving into the implementation of gesture control technology within healthcare environments. Our primary objective centers on addressing the distinctive challenges encountered by healthcare professionals, with a key focus on providing an intuitive and hands-free method for interacting with the physical surroundings. This research seeks to explore the potential of gesture control as a catalyst for revolutionary changes in healthcare practices, presenting functionalities that encompass seamless control over air conditioning, fans, gates, lighting, and doors.

The significance of this research transcends mere technological convenience, delving deeply into broader implications for hygiene and accessibility within healthcare facilities. In response to the escalating emphasis on infection control and the imperative to minimize physical contact, the adoption of gesture control emerges as a promising avenue to alleviate these concerns. By empowering healthcare professionals to execute tasks like adjusting environmental conditions or accessing secured areas through straightforward, non-contact gestures, this research aspires to make a substantive contribution to cultivating a safer and more streamlined healthcare environment.

The intricate scope of this inquiry encompasses not only the technical nuances associated with implementing gesture recognition systems but also the pragmatic implications for healthcare workflow. As we navigate the

Grenze ID: 01.GIJET.10.2.531_5

development of a comprehensive system tailored to the unique needs of healthcare professionals, our anticipations extend beyond shedding light on the capabilities of gesture control in healthcare. We foresee that the outcomes of this research will act as a catalyst, paving the way for future advancements in technology applications that are not only user-friendly but also uphold stringent hygiene standards within medical settings.
The remaining sections of the study are structured as follows: Section II presents the literature review; Section III elaborates on the suggested system; Section IV provides an explanation of the findings; and Section V wraps up the work.

II. LITERATURE REVIEW

A framework for the recognition of Indian sign language was created by Anand et al. [1]. It is believed that the two main parts of the proposed system, feature extraction and classification, include a pattern recognition method. To recognize sign language, Discrete Wavelet Transform (DWT) based feature extraction and the closest neighbour classifier work together. The recommended hand gesture recognition system uses a cosine distance classifier to achieve a maximum classification accuracy of 99.23%, in accordance with the experimental findings.
A PCA-based system for hand gesture recognition was created by Ahuja et al. [2]. The authors of this paper developed an effective template matching system that may be used in human robotics and related domains, as well as a database-driven method for hand gesture identification based on thresholding and skin colour model approaches. The hand region is then divided into segments using the YCbCr colour space and a skin colour model. Thresholding is used in the following step to distinguish between the foreground and background.
Cheok et al. [3] reviewed a variety of methods that have been previously proposed by various academics for hand gesture and sign language recognition. Sign language is the only means of communication for the deaf and dumb. These individuals with physical impairments communicate their feelings and ideas to others using sign language. They claim that hand gesture recognition technologies enable human-computer interaction. The two main applications they have employed are recognition of sign language and gesture-based control. Gestures are movements of the hand, finger, arm, head, face, and body that indicate message. It is believed that a hand gesture recognition system is a more powerful and intuitive tool for computer-human interaction. Applications encompass virtual prototyping, sign language analysis, and medical education.
A method for displaying continuous, real-time sign language gesture recognition was proposed by Lai et al. [4]. Starting with the gesture input, the statistical analysis is based on posture, location, orientation, and motion. Sign language is the most expressive medium for the hard of hearing. A recognizer must be able to instantly identify continuous sign vocabularies. To find the end points in a gesture input sequence, they applied the End-Point problem.
A reliable technique for hand gesture detection based on convolutional neural networks was presented by Yingxin et al. [5]. This method uses the network to automatically extract the spatial and semantic features of the hand motion. The combined effects of this method's data pre-processing and improved Convolutional Neural Network topology improve hand gesture detection performance. The suggested approach is competitive and successful, as demonstrated by the experimental results on the self-constructed dataset and the Cambridge Hand Gesture Dataset.
Using computer vision and image processing techniques, Dehankar et al. [6] suggested a novel Real Time End Point Identification Method from a real-time video of a hand motion. Accurate End Point Identification (AEPI), on which the Real Time End Point Identification (RTEPI) approach is based, allows the AEPI method to operate on hand motions recorded in real time using a laptop or web camera. The AEPI method, which determines the accurate method, receives the proper frame for real-time processing from the RTEPI method. The AEPI technique has been employed to tackle issues related to changing backgrounds, brightness, blurring, and so forth.

III. PROPOSED METHOD

The main focus of study is creating an effective gesture detection system that can quickly and reliably identify and assess hand and body movements. The use of deep learning methods, such as recurrent and convolutional neural networks, is frequently credited with the improvement in this field.
This study employs OpenCV, a Python library, to manage real-time data, specifically for the recognition and processing of hand gestures. Figure 1 illustrates the progression of data from one phase to another. The following common steps are included in the suggested machine learning approach for hand gesture recognition:
1. Data Collection and Pre-processing: Put together a varied collection of hand gesture photos or films that were taken from different perspectives, with different lighting and backgrounds. To guarantee consistent input data, pre-process the photos/videos by resizing, normalising, and removing noise.

2. Feature Extraction: Derive meaningful features from the pre-processed images/videos. This can be achieved through traditional computer vision techniques such as histogram of oriented gradients (HOG), scale-invariant feature transform (SIFT), or more contemporary techniques like Convolutional Neural Networks (CNNs). Additionally, a pre-trained CNN model may be utilized for feature extraction.
3. Data Augmentation: Apply data augmentation techniques for model generalization. These techniques involve rotation, scaling, flipping, and adding noise to artificially enhance the size and diversity of the training dataset.
4. Model Selection: Choose a suitable machine learning algorithm or model for gesture recognition.
5. Model Training: Divide the dataset into training, validation, and test sets. Conduct fine-tuning and hyper parameter tuning by experimenting with various model architectures, hyper parameters, and optimization strategies such as learning rate scheduling, dropout, and batch normalization.
6. Evaluation: Assess the model's performance on the test set using metrics like accuracy, precision, recall, F1-score, and the confusion matrix.

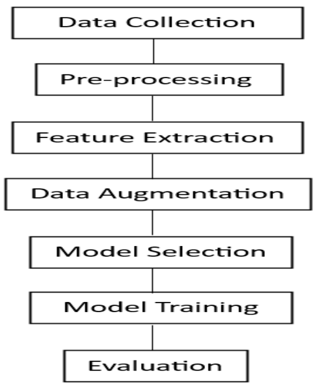

Figure 1: Hand gesture recognition system

A. Data collection: The hand gesture recognition system can be fed raw data using basically three different approaches. The first approach, known as the glove-based approach, uses a colour glove or data glove. In the data glove approach, the primary drawbacks include its high cost due to sensor nodes and a lack of naturalness caused by the need to wear the data glove consistently. However, colour gloves offer a more economical alternative, as they do not incorporate sensors either inside or outside the gloves. This method proves to be a robust solution for hand gesture recognition systems. The second approach utilizes one or more cameras to collect raw input for hand gestures, known as the vision-based approach. This method's strength lies in its natural and convenient communication, but it is susceptible to interference from complex backgrounds. The last approach is a hybrid method that combines the two techniques to collect raw hand data.
B. Segmentation techniques: Effectively segmenting the hand is a crucial aspect of human hand gesture recognition systems, as the segmentation process is significantly impacted by factors such as diverse illumination conditions, static clutter backgrounds, dynamic backgrounds, and the presence of multiple gestures in the same background. Inaccurate segmentation can result in data loss, preventing the correct input of gestures into the recognition system and leading to incorrect recognition. The subsequent stages heavily rely on the accuracy of hand segmentation. Several commonly used hand segmentation techniques are outlined below:
1. Skin Colour Segmentation: Skin colour segmentation is frequently employed in hand gesture segmentation procedures. Researchers commonly utilize colour spaces such as RGB, YC_bC_r, and HSV. The YCbCr and HSV colour spaces show superior clustering properties than the RGB colour system, despite the RGB colour space combining three types of colour information and producing a significant correlation between skin range and luminance. YC_bC_r, obtained through linear transformation of RGB primary colour, proves to be a discrete space with improved clustering results and computational efficiency. The YC_bC_r colour model is particularly well-suited for human skin colour segmentation, although HSV colour space, being continuous, may misclassify if the background contains objects of similar colour to the skin.
2. Background Subtraction: Background subtraction is a straightforward frame differencing method. This approach defines the moving target as the difference between the current image and the background image. Its strengths include speed, ease of implementation, effective detection, and the provision of complete feature data for the target. However, this method is sensitive to dynamic changes in factors such as indoor lighting.

3. Contour Matching: Contour matching involves the sequence of pixels representing the line or curve of an object in an image. Widely used for image object classification and recognition, contour matching is preferred over template matching because it extracts only the edge points rather than the entire image. This method offers advantages such as sharp matches, speed, and invariance under imaging transformations like rotation, intensity scaling, and translation. However, its weakness lies in the involvement of a small number of pixels, potentially hindering recognition due to the limited influence of each pixel.

C. Hand gesture recognition techniques: Features are used as input during the recognition phase to create class-labelled outputs that represent the intended motions. The right choice of gesture parameters and attributes has a big impact on recognition accuracy. The most popular methods for recognizing hand gestures are as follows:

1. Dynamic Time Warping (DTW): DTW addresses both static and dynamic aspects of gestures. It serves a dual purpose by identifying candidate gestures and validating their similarity. The DTW algorithm compares the similarity between the captured gesture and those in a pre-trained database, offering the advantage of time compression.

2. Artificial Neural Network (ANN): Numerous researchers have explored gesture recognition using neural networks as classifiers. A multilayer feedforward network, configured with one input layer, two hidden layers, and an output layer (e.g., 32-100-60-9), has been employed. The widely used neural network algorithm is the feedforward backpropagation algorithm in pattern recognition. However, a weakness of artificial neural networks is their susceptibility to time delays.

3. Support Vector Machine (SVM): Support Vector Machine is a popular algorithm frequently utilized by researchers as a classifier. It is a supervised learning method applicable to both recognition and regression challenges. SVM performs well when there is a clear margin of separation. However, its weakness lies in requiring higher training time for large datasets and when there is more noise in the dataset. A bag-of-features is provided as an input vector for a multiclass SVM to classify different types of gestures.

IV. EXPERIMENTAL RESULTS

Gesture recognition plays a pivotal role in accurately detecting and analysing human movements, offering applications such as device control, virtual environment interaction, and feedback for physical therapy and rehabilitation exercises. The output of gesture recognition can take the form of commands to manipulate virtual objects or visual representations of movements, such as graphs or animations. The overarching goal is to facilitate seamless and instinctive interactions between humans and machines, enhancing user experiences and improving task efficiency and accuracy.

In recent years, advancements in computer vision, machine learning (ML), and sensor technology have significantly boosted gesture recognition technology. Modern systems, armed with more precise algorithms, can effectively detect and interpret diverse human movements, including hand gestures, body gestures, and facial expressions. The recognition accuracies have seen substantial improvement, with traditional computer vision techniques achieving 80% to 90% accuracy, deep learning methods like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) achieving 90% to 99% accuracy, and more recent Transformer-based models showing even better results, albeit with a demand for substantial amounts of data. The proposed method falls within the range of 95% to 97% accuracy, with variability influenced by numerous hyperparameters.

Gesture recognition has found applications in virtual and augmented reality, enabling users to interact with digital objects through natural movements. In healthcare and rehabilitation, technology provides real-time feedback on movement precision during therapy exercises. In industrial automation and robotics, gesture recognition is utilized for machinery control, resulting in enhanced efficiency and productivity.

The output of gesture recognition varies based on the application. In virtual reality, it might involve visual representations of hand movements for user interaction with virtual environments. In healthcare, the output could be a score or visual feedback indicating the accuracy of a patient's movement or exercise. In industrial automation, the output may include commands for robots or machines to execute specific actions.

As gesture recognition technology continues to evolve, it has the potential to revolutionize human-machine interactions, offering users a more natural and intuitive experience. The ongoing advancements in technology suggest that we can expect even more sophisticated and precise gesture recognition systems, opening up new possibilities for various applications and use cases in the future.

The model performed well based on the plot and it even score 100% accuracy on the 4th epoch. This could probably because the images are too simple and dataset are not enough. The experiment workflow is repeated 5 times and additional images are added in each repetition. The accuracy and loss curve of the model on the 5th iteration with 2200 images is shown in Figure 3.

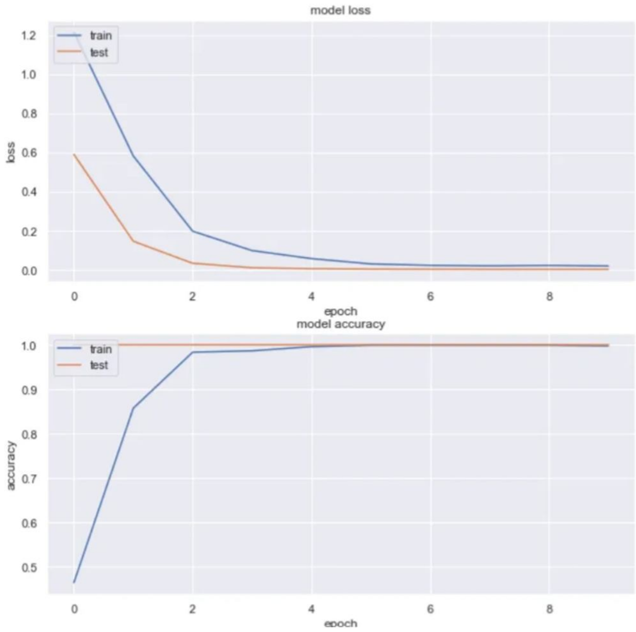

Figure 2 shows both loss curve and accuracy curve of the proposed model at 4th epoch.

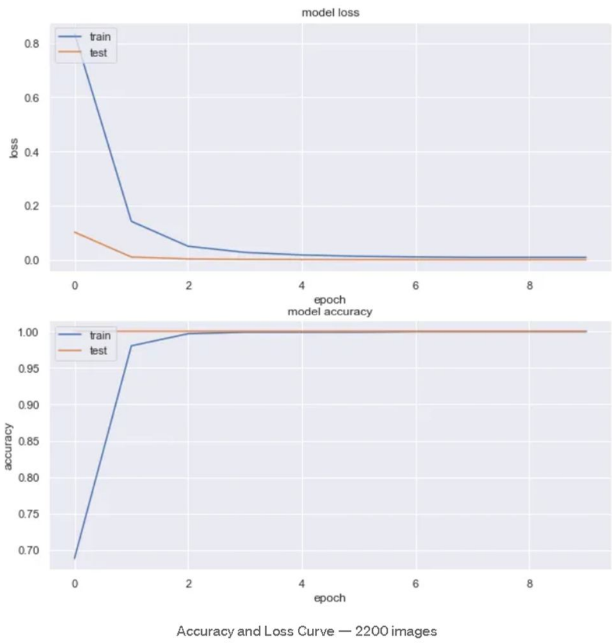

Figure 3: Loss curve and Accuracy curve at 5th epoch

V. CONCLUSION

Utilizing Machine Learning (ML) in a Gesture Handler system has shown to be quite beneficial, especially in the context of developing technologies. An intelligently conceived and built hand gesture detection system running as a stand-alone application can control a variety of UI components and applications, including the VLC media player. Additionally, this system can intermittently monitor and validate the accuracy of the gesture recognition process by ensuring that the shapes of layout pictures and individual pictures align. Such an advanced system provides valuable support in seamlessly controlling user interfaces and applications through intuitive hand gestures. In summary, gesture recognition technology has experienced significant advancements, driven by breakthroughs in computer vision, ML, and sensor technology. Its broad array of applications, spanning virtual and augmented reality, healthcare, and industrial automation, has the potential to transform human-machine interactions. By offering a natural and intuitive mode of communication, gesture recognition technology holds the promise of enhancing user experiences, improving efficiency, and boosting accuracy across various domains. As research and development progress, we can anticipate the emergence of even more sophisticated and precise gesture recognition systems in the future, unlocking fresh opportunities for innovative applications and use cases.

FUTURE WORKS

Moving forward, future research endeavors in this domain should explore avenues for refining existing models and advancing the state-of-the-art in PD detection. Investigating ensemble methods that combine the strengths of multiple algorithms could potentially yield enhanced diagnostic accuracy by leveraging the diverse capabilities of each model. Furthermore, the exploration of advanced feature engineering techniques and the integration of deep learning architectures may unlock additional discriminatory power in capturing complex disease patterns. Longitudinal studies with larger and more diverse datasets are essential to validate and generalize the observed trends, ensuring the robustness of the developed models.

Moreover, the integration of multimodal data sources, such as combining clinical and imaging data, holds promise for a more holistic understanding of PD. Real-time applications of these models in clinical settings could revolutionize early diagnosis and intervention. Continuous collaboration between machine learning researchers, clinicians, and data scientists is imperative for translating these findings into practical, clinically applicable tools. The interdisciplinary nature of future work will be pivotal in developing reliable and interpretable machine learning models that significantly contribute to the field of PD diagnosis, ultimately improving patient outcomes and healthcare practices.

REFERENCES

[1] Mathavan Suresh Anand, Nagarajan Mohan Kumar and Angappan Kumaresan, “An Efficient Framework for Indian Sign Language Recognition Using Wavelet Transform,” Circuits and Systems, vol. 7, pp 1874- 1883, 2016.

[2] Mandeep Kaur Ahuja, and Amardeep Singh, “Hand Gesture Recognition Using PCA,” International Journal of Computer Science Engineering and Technology (IJCSET), vol. 5, issue 7, pp. 267-27, July 2015.

[3] Ming Jin Cheok, Zaid Omar, and Mohamed Hisham Jaward, “A review of hand gesture and sign language recognition techniques,” International Journal of Machine Learning and Cybernetics, vol. 10, pp. 131-153, August 2017.

[4] Hsiang Yueh Lai, Hao Yuan Ke, and Yu-Chun Hsu, “Real-time Hand Gesture Recognition System and Application,” Sensors and Materials, vol. 30, issue 4, pp. 869-884, 2018.

[5] Xing Yingxin, Li Jinghua, Wang Lichun, and Kong Dehui, “A Robust Hand Gesture Recognition Method via Convlutional Neural Network,” In proceedings of 6th IEEE International Conference on Digital Home (ICDH), pp. 64-67, 2016.

[6] A V Dehankar, Sanjeev Jain, and V M Thakare, “Performance analysis of RTEPI method for real-time hand gesture recognition,” In proceedings of IEEE International Conference on Intelligent Sustainable Systems (ICISS), pp. 1031-1036, 2017.