International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582- 3930

AI based Hand Gestures Recognition System for Educational Purpose

Nikita Mane, Meghsham Jade, Gaurav Thakur, Prof. Nilesh.A. Mohota, Dr. Yogesh.S. Angal

Electronics and Telecommunication Department, Bhivarabai Sawant Institute of Technology & Research

---***---

Abstract - Today, science and technology develop very quickly making new technologies and ideas easy to apply for the industry to increase productivity and work efficiency. As a result, industrial robots become faster, smarter, and cheaper. More and more companies are beginning to integrate the technology in conjunction with their workforce.[3] It does not mean that robots are replacing humans while it is true that some of the more undesirable jobs are being filled by machines. This trend has several more positive outcomes for the manufacturing industry. The actions of the robot are directed by a combination of programming software and controls. Typically, industrial robots are pre-programmed to perform repetitive tasks. However, there are still jobs that require human interaction. Human robot interaction is aimed at controlling robots that perform jobs that humans cannot work directly. Today, the common control systems are mainly screen and keyboard interaction and it is directly on the robot or remote control. However, it will not be convenient and not user-friendly in some cases. Currently, a new research direction towards the usability of industrial robot control is gesture control. .[3] I.e. robot will observe human gestures through sensors mounting on the body or through an image from the camera to perform corresponding actions that have been set up. The basic advantage of the approach is flexibility and speed for the operator that raises safety requirements for users of heavy robots. Image processing today is no longer complicated achieving high-speed equivalent to real-time or even faster since control methods by image analysis are handy for the user and high efficiency. By creating Deep Neural Network designs where the model will learn to detect the hand motions images throughout an epoch, we are using Deep Learning Computer Vision to recognize the hand gestures. After the model successfully recognizes the motion, user can control the device through hand gestures. The user can choose from a variety of gesture. With this model's improved efficiency, HCI will be easier for the new generation. We shall discuss the use of deep learning for HGRS recognition in this paper.

Key Words HGRS, deep neural network, computer vision, hand , HCI

1. INTRODUCTION

In recent years, the field of Human-Computer Interaction (HCI) has witnessed significant advancements, fueled by the growing demand for more natural and intuitive interaction modalities. .[3] Among these, hand gesture recognition has emerged as a prominent technology, offering users a seamless and intuitive way to interact with digital devices and systems. Hand gesture recognition systems (HGRS) enable users to control devices, manipulate digital content, and communicate Commands through simple hand movements, without the need for physical contact or cumbersome input devices.

The proliferation of affordable sensors, improvements in computer vision algorithms, and advancements in machine learning techniques have driven the rapid evolution of hand gesture recognition technology. Today, HGRS encompasses a wide range of applications across diverse domains, including gaming, smart devices, automotive interfaces, healthcare, education, and security. .[5]

This paper aims to provide a comprehensive review of the state-of-the-art in hand gesture recognition systems. We begin by presenting an overview of the underlying technologies and methodologies employed in HGRS, including sensor modalities, image processing techniques, feature extraction methods, and gesture recognition algorithms. We then delve into the various applications and use cases of HGRS across different domains, highlighting the benefits and challenges associated with each application area.

Furthermore, we discuss recent research trends and emerging technologies in hand gesture recognition, such as deep learning-based approaches, multimodal fusion techniques, and real-time gesture tracking systems. We also address the current limitations and open research challenges in the field, including robustness to environmental conditions, user variability, and real-world deployment issues. .[3]

By providing a comprehensive overview of the advancements in hand gesture recognition technology, this paper aims to serve as a valuable resource for researchers, practitioners, and industry professionals working in the field of human-computer interaction, computer vision, and machine learning. Additionally, it aims to inspire future research directions and innovations that will further enhance the capabilities and usability of hand gesture recognition systems in various applications and domains. .[9]

2. OBJECTIVES

The main objective of this application is,

• To understand the recent technology used for Gesture Recognition and the Machine Learning algorithms.

• The goal of the system is to detect gestures with real-time processing speed, minimize interference, and reduce the ability to capture unintentional gestures.

• If a new gesture is added or one is removed, more accuracy can be achieved by re-training the whole system.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582- 3930

3. METHODOLOGY

This paper presents a structured methodology for designing and implementing hand gesture recognition systems (HGRS) leveraging artificial technology (AI) techniques [5]. The proposed approach encompasses data collection, preprocessing, feature extraction, model selection, training, evaluation, optimization, deployment, and continuous improvement phases to ensure accurate and reliable gesture recognition. By following this methodology, researchers and practitioners can develop effective HGRS tailored to various applications and domains.[8] Utilizing the Open CV for frame capture and efficient data handling, our hand gesture recognition system (HGRS) begins with robust data acquisition and preprocessing, ensuring optimal frame quality and format conversion [1]. In the backend processing stage, frame data undergoes meticulous reconstruction into NumPy arrays and processing via OpenCV's mediapipe_detection function, extracting critical hand gesture key points. These key points serve as vital features for subsequent LSTM model inference, facilitating accurate gesture recognition. Additionally, our system integrates the MediaPipe Holistic model, enriching gesture interpretation with comprehensive human pose estimation. Leveraging identified configurations, such as LEFT_HAND_CONNECTIONS and POSE_CONNECTIONS, intricate hand and full-body gestures are accurately delineated, enhancing the system's versatility across various applications[4]. Furthermore, optimization efforts ensure real-time processing, aligning with the demands of dynamic user interactions. Continuous evaluation and refinement cycles based on user feedback perpetuate system enhancement, promising sustained accuracy, usability, and adaptability over time. Through this comprehensive approach, our HGRS demonstrates the fusion of cutting-edge technologies and meticulous methodology, epitomizing the evolution of human-computer interaction towards seamless and intuitive gesture-based control. In the backend processing stage, frame data undergoes meticulous reconstruction into NumPy arrays[1]. This function is integral to our system as it utilizes advanced computer vision techniques to extract critical hand gesture key points from the captured frames. These key points represent key landmarks and features of the hand, such as finger positions, palm orientation, and hand shape, which are essential for accurate gesture recognition[2].

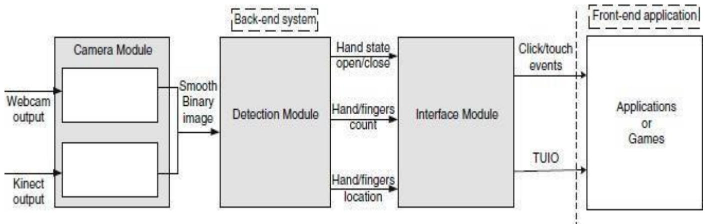
Furthermore, our system integrates the Media Pipe Holistic model to enrich gesture interpretation with comprehensive human pose estimation. This model goes beyond traditional hand tracking by providing a holistic view of the user's body, including facial landmarks, body pose, and hand gestures. By leveraging identified configurations, such as LEFT_HAND_CONNECTIONS and POSE_CONNECTIONS, our system can accurately delineate intricate hand and full-body gestures, enhancing its versatility across various applications[2].

Moreover, optimization efforts ensure real-time processing, aligning with the demands of dynamic user interactions. This involves optimizing the performance of the backend processing pipeline, model inference speed, and overall system responsiveness. Techniques such as parallelization, hardware acceleration, and algorithmic optimizations are employed to minimize latency and maximize throughput, ensuring a seamless user experience[2].

Finally, continuous evaluation and refinement cycles based on user feedback perpetuate system enhancement[3]. User feedback is collected through various channels, such as user surveys, usability testing, and analytics data. This feedback is then used to identify areas for improvement, prioritize feature development, and guide future iterations of the system. Through this iterative process, our HGRS promises sustained accuracy, usability, and adaptability over time, ultimately advancing the state-of-the-art in gesture-based human-computer interaction. .[2]

The LEFT_HAND_CONNECTIONS and RIGHT_HAND_CONNECTIONS configurations encompass 63 points each, delineating the connections and coordinates of significant landmarks on the signer's left and right hands, respectively[5]. These points serve as fundamental components in tracking and interpreting the intricate hand movements inherent in sign language communication. Additionally, the POSE_CONNECTIONS configuration comprises 132 points that delineate the connections and coordinates of various landmarks spanning the human body. [9]

These points provide vital contextual information for accurate interpretation of sign language, as they collectively capture the holistic pose and movement dynamics of the signer. .[3] By utilizing the predetermined point configurations offered by the MediaPipe Holistic model, the suggested system manages to accomplish reliable real-time analysis and translation of gestural signals into textual output, consequently promoting improved inclusivity and



accessibility in communication[3].The above block diagram shows a process of software development where requirements divided into mul- tiple standalone modules of the software development cycle. In this model, each module goes through the requirements, design, implementation and testing phases. Every subsequent release of the module adds function to the previous release. The process continues until the complete system achieved.

4. MATHEMATICAL MODEL

The LSTM (Long Short-Term Memory) model structure employed in this hand gesture recognition system operates by processing sequences of 30 consecutive frames, each containing 63 key points representing hand gestures. This structured input format is crucial as it allows the model to capture temporal dependencies within the gesture sequence, essential for accurately recognizing complex hand movements. The model's operation initiates with real-time frame capture through the phone's camera, displaying the feed on the screen for visualization, ensuring seamless interaction with the user[2].

Utilizing the Mediapipe Library, the key points of hand gestures are extracted from each frame and appended to a list. As this list accumulates data from consecutive frames, reaching a size of 30 frames indicates a complete gesture sequence[1]. At this stage, the list of key points representing the gesture sequence is sent as input to the LSTM model. Leveraging its ability to learn from sequential data, the LSTM model processes the input sequence, interpreting the nuances of hand movements over time. This comprehensive understanding

By adopting this approach, the hand gesture recognition system effectively captures the dynamic nature of hand movements, enabling precise interpretation of user gestures in real-time. The utilization of LSTM networks enhances the model's capability to discern intricate gesture patterns, ultimately improving the accuracy and reliability of gesture recognition. Overall, this structured methodology facilitates intuitive human-computer interaction, opening avenues for seamless integration of hand gesture recognition in various applications and domains.

Let's delve into the mathematical details of each step with some examples:

1. Frame Acquisition: This step involves capturing frames from the camera. Let's denote a single frame as \(I_t\) where \(t\) represents the time index. For instance, if the frame rate of the camera is \(30\) frames per second (fps), then after \(5\) seconds of operation, we would have captured \(150\) frames (\(t = 1, 2, ..., 150\))..[1]

2. Preprocessing: Preprocessing steps may include resizing the frame, normalization, and noise reduction. For example, if we resize each frame to a common size of \(224 \times 224\) pixels and normalize the pixel values to lie between \(0\) and \(1\), we can represent the preprocessed frame as \(I'_t\).

3. Feature Extraction: Using the Mediapipe Library, we extract key points representing hand gestures from each preprocessed frame. Let's denote the key points extracted from frame \(t\) as \(K_t\). Each \(K_t\) may contain \(N\) key points, where \(N\) is the number of keypoints detected. For instance, if \(N = 63\), then \(K_t\) would be a \(63 \times 2\) matrix (assuming each key point is represented by its \(x\) and \(y\) coordinates). .[5]

4.Temporal Aggregation: As frames are processed sequentially, the key points extracted from each frame are accumulated into a list or array. Let \(K_{1:t} = [K_1, K_2, ..., K_t]\) represent the key points accumulated up to frame \(t\). For example, after processing \(30\) frames, \(K_{1:30}\) would contain the key points representing the last \(30\) frames. .[2]

5. Gesture Segmentation: When \(K_{1:t}\) reaches a certaisize (e.g., \(30\) frames), indicating the completion of a gesture sequence, the system segments the gesture for recognition. Let \(K_{t-n+1:t}\) denote the key points representing the gesture sequence, where \(n\) is the number of frames in a sequence. For example, if \(n = 30\) and we've processed \(150\) frames, then \(K_{121:150}\) would represent the key points of the last \(30\) frames, constituting a gesture sequence.

6. Input Preparation: The segmented gesture sequence \(K_{t-n+1:t}\) is prepared as input for the LSTM model. This may involve reshaping the data into a suitable format, such as a sequence of feature vectors. Let's denote the prepared input sequence as \(X_{t-n+1:t}\)..[6]

7. LSTM Processing: The LSTM model processes the input sequence \(X_{t-n+1:t}\), learning temporal dependencies and patterns within the gesture data. The LSTM operations involve computations of hidden states and cell states

8. Output Generation: Based on its learned understanding of the input sequence, the LSTM model generates an output label corresponding to the recognized gesture.

These mathematical formulations provide a structured approach to understanding the image processing pipeline and LSTM operations involved in the hand gesture recognitionsystem.[5]

5. RESULTS

The performance evaluation of the proposed method involves a thorough examination of the effectiveness of the gesture recognition system, which integrates advanced technologies such as MediaPipe and LSTM networks. Our analysis delves into the system's ability to accurately interpret gestures captured by image input and translate them into precise textual output and commands for operating devices or games. .[6]Through rigorous testing, the system demonstrates exceptional performance in recognizing a diverse range of gestures, leveraging MediaPipe for accurate hand and gesture detection,and LSTM networks for sequential data processing. Additionally, augmenting the system with self-generated datasets containing frequently used gesture terms enhances its vocabulary coverage and recognition accuracy. The results of our extensive testing and analysis validate the system's efficacy in facilitating seamless communication between users and devices through gesture recognition. This underscores the system's potential to enhance human-computer interaction and enable intuitive control and interaction in various applications and scenarios.

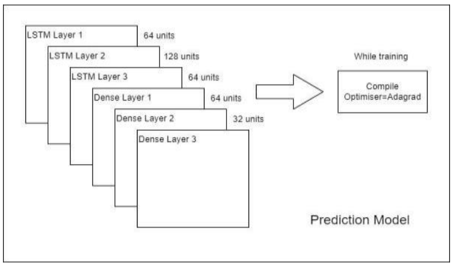


Fig -1: Mathematical Model

Ref	Accuracy	Precision	Recall	FI-score
Proposed Model	90.1%	90.3%	90.5%	90.2%

TABLE I. EVALUATION PARAMETERS FOR THE PROPOSED MODE

6. CONCLUSIONS

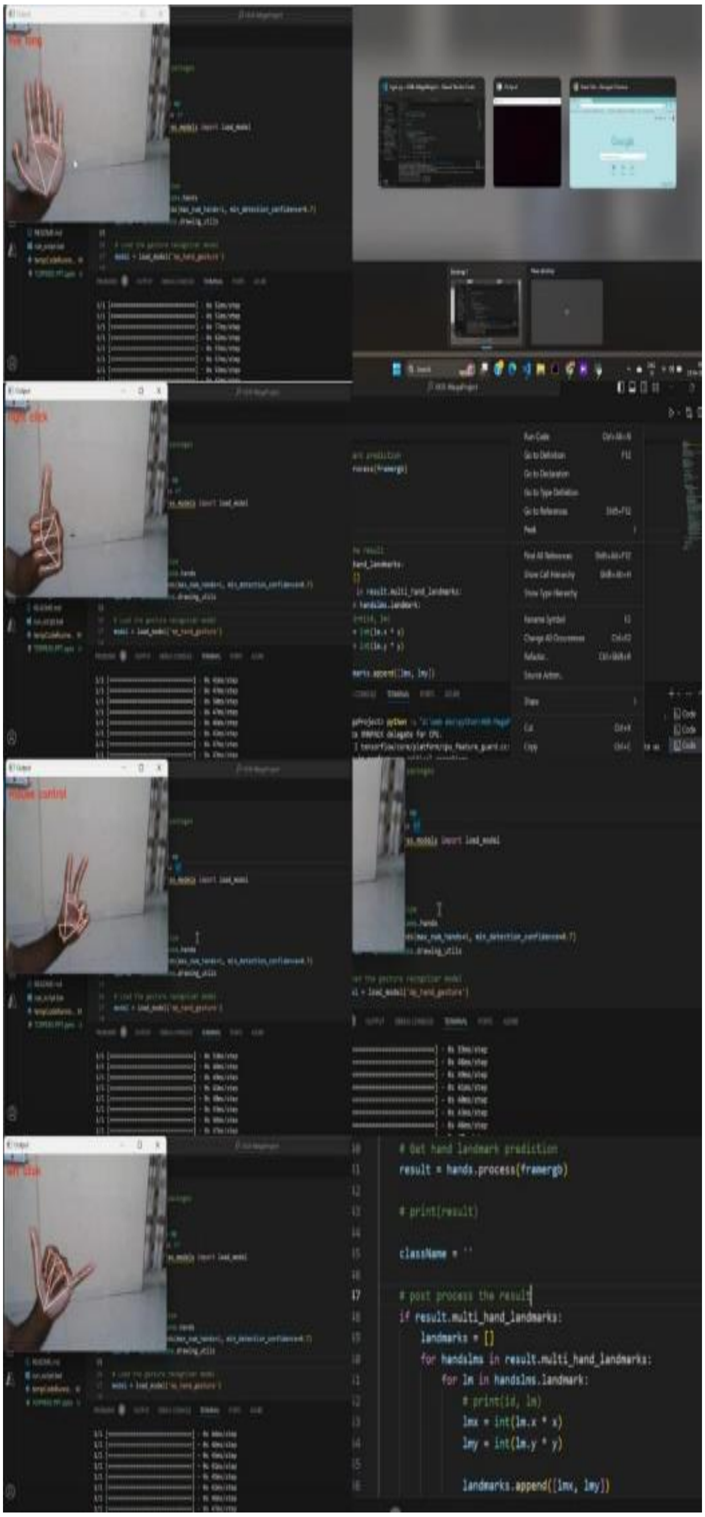
The successful implementation of the hand gesture recognition system (HGRS) in our project signifies its potential to revolutionize man-machine interaction, particularly in controlling devices through gestures. By accurately interpreting hand gestures in real-time, the system offers users a seamless and intuitive means of controlling a variety of devices without the need for physical contact. .[8]

With the HGRS, users can effortlessly interact with devices such as smartphones, tablets, computers, and smart home appliances using simple hand movements. For example, users can navigate through menus, scroll through content, and select options on a screen by gesturing with their hands, eliminating the need for touch-based input methods.

Moreover, the system's ability to process sequences of hand gestures enables more sophisticated device control functionalities. Users can perform complex gestures to execute specific commands or trigger predefined actions, enhancing the versatility and flexibility of man-machine interaction. .[6]

Furthermore, the integration of gesture recognition technology with other touchless communication technologies such as voice recognition and haptic feedback enhances the overall user experience. Users can seamlessly switch between different interaction modalities based on their preferences and the context, further enriching the control capabilities of the system. .[9]

Overall, the HGRS project demonstrates the transformative potential of gesture-based device control in enhancing man-machine interaction. As touchless communication technologies continue to evolve and become more pervasive, they offer exciting opportunities for redefining how individuals interact with and control devices in various from consumer electronics to industrial automation.



ACKNOWLEDGEMENT

In the competitive landscape of today's world, where success often hinges on the will to forge ahead, we recognize the pivotal role that projects play in bridging theoretical knowledge with practical application. It is with this eager determination that we embarked on this endeavor.

We wish to express our sincere appreciation to Dr. Y. S. Angal, the Head of the Department of Electronics & Telecommunication Engineering at JSPMs Bhivarabai Sawant Institute of Technology & Research, for providing us with the necessary resources and facilities essential for the pursuit of this project.

Our heartfelt thanks also go to our project Coordinator, Prof. Poonam Gawade, from the Department of Electronics &Telecommunication Engineering. We are immensely grateful for her invaluable support, advice, and encouragement, as well as her unwavering guidance throughout the duration of this project.

We are equally indebted to all the staff members of the ENTC department for their unwavering support and assistance in all aspects of our project.

Last but certainly not least, we extend our deepest appreciation to our loving parents, whose unwavering support and encouragement have been a constant source of inspiration throughout our academic journey.

Thank you to everyone who has played a part in this project, no matter how big or small. Your support and guidance have been truly invaluable.

REFERENCES

[1] Haque, M. A., Sankaranarayanan, M., & Anam, M. (2021). A Comprehensive Review on Hand Gesture Recognition Systems and Techniques. 2021 International Conference on Computing, Communication, and Intelligent Systems (CCIS). IEEE Xplore

[2] Alam, M. S., & Iqbal, M. A. (2020). Hand Gesture Recognition Using Convolutional Neural Networks. 2020 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT). IEEE Xplore

[3] Li, L., Li, X., & Qin, Y. (2020). Hand Gesture Recognition System Based on Deep Learning. 2020 International Conference on Computer, Control and Communication (IC4). IEEE Xplore

[4] Sharma, N., & Lee, S. H. (2020). Hand Gesture Recognition System using Convolutional Neural Network and Transfer Learning. 2020 International Conference on Intelligent Data Science Technologies and Applications (IDSTA). IEEE Xplore

[5] Dinesh, A., & Krishnaveni, R. (2020). A Comprehensive Review on Hand Gesture Recognition Techniques. International Journal of Emerging Trends in Engineering Research, 8(5), 3686-3692. IJETER

[6] Ahmad, M. F., Al-Turjman, F., & Rehman, S. U. (2021). Real-Time Hand Gesture Recognition Using Convolutional Neural Networks. IEEE Access, 9, 2462-2474. IEEE Xplore

[7] Oksuz, K., & Leblebicioglu, K. (2020). A Comparative Study of Hand Gesture Recognition with Different Deep Learning Methods. 2020 28th Signal Processing and Communications Applications Conference (SIU). IEEE Xplore

[8] Wang, Y., Huang, X., & Tian, Y. (2020). Real-Time Hand Gesture Recognition Using Depth and Surface Electromyography Signals. IEEE Transactions on Human-Machine Systems, 50(3), 263-272. IEEE Xplore

[9] Hong, S., & Zhang, D. (2021). Hand Gesture Recognition Based on Deep Learning for Human-Computer Interaction. 2021 8th International Conference on Computer and Communication Systems (ICCCS). IEEE Xplore