

System Control by Hand Gestures using Artificial Intelligence

Shruthi A ¹, Ganga B M ², Aswini S³ & Abhinandan B S ⁴

¹ Assistant Professor, ²Associate Professor, ³Associate Professor
CSE, ACSCE, Bengaluru

⁴Student, CSE, ACSCE, Bengaluru

DOI: <https://doi.org/10.34293/acsjse.v2i2.34>

Abstract - —"Information era" is what we call the present world and there are many advancements in every field across the globe, especially in the field of AI computer vision. Researchers have found interesting results experimenting upon the fields which has lead to creation of many other interesting fields such as object detection, object tracking, image processing and more. Here in our project keeping "Vision Based Human Detection Techniques: A descriptive review", by Shahriar Shakir Sumit, Dayang rohaya among Rambli and Seyedali Mirjalili as the base paper, were implementing a project of hand tracking and gesture recognition to combine AI with computer vision to achieve complete control over computer without the use of Keyboard or mouse. The project involves simple and predefined gestures which is recognized by AI via camera in the computer and performs tasks accordingly. This project can be helpful to keep a wide step ahead to overcome the increasing problem of e-waste as the project completely eliminates the use of keyboards and mouse resulting in less production and hence less e- waste.

I INTRODUCTION

In cutting-edge international, era is developing very fast, and we have become in contact with special new technology day with the aid of using day. Here, one of the booming technology of pc technological know-how is Artificial Intelligence which is prepared to create a brand-new revolution withinside the international with the aid of using making shrewd machines. The Artificial Intelligence is now all round us. It is presently running with loads of subfields, starting from fashionable to specific, along with self-using cars, gambling chess, proving theorems, gambling music, Painting, etc. Following are a few foremost motives to find out about AI: With the assist of AI, you could create such software program or gadgets which could resolve actual international issues very without problems and with accuracy along with fitness issues, marketing, site visitor's issues, etc. With the assist of AI, Along with Siri, Cortana, Google Assistant, and other digital assistants, you may develop your own. You could create such Robots with the aid of AI.

This might be effective in a situation where people's lives are in danger. AI paves the way for a variety of new technologies, devices, and opportunities. However, we're the usage of synthetic intelligence to increase a software program which solves one such actual international issues, e-waste. Annually 53.6 million heaps of e-waste is produced globally i.e. 7. three kg in line with individual and round 7.seventy one million heaps of e-waste is produced on my own with the aid of using India and is growing on the charge of 8.86% in line with year. Major a part of this e-waste are broken digital gadgets, gadgets that are antique and which cannot be up to date any further, keyboards, mouse, video display units etc. With our undertaking we need to take away the usage of keyboard and mouse via which we will

lessen the quantity of e-waste produced to positive extent. For this we researched many strategies and opportunities however now no longer all current era have been useful to resolve the trouble via our perspective, aside from one approach that's the usage of mediapipe framework via which we have been capable of seize information or capabilities required from stay circulate or stay footage.

II EXISTING SYSTEM

There is no published research on this specific topic, according to our literature review. However, it appears that there are only a small number of texts in circulation that examine the use of more conventional approaches, such as the usage of You Look Only Once (YOLO) and You Look Only Once Version 6 (YOLO V6) which also have other shortcomings such as these methodologies can be best suited for still images or if the video footage is fed in breaks as inputs and these shortcomings must be addressed.

For the above reasons, on this paper we advocate the use of Mediapipe framework developed by Google for sole purpose of providing solutions for artificial intelligence and gesture/ moment detection problems.

III PROPOSED SYSTEM

In this we successfully carried out the procedural process of implementing virtual keyboard and virtual mouse and adding functionalities to it, but the major challenging part was to track the moments of the user or users hands to be specific and capture the coordinates so as it use it as input for the project.

Though there were many traditional methodologies available like You Look Only Once (YOLO), due to its many disadvantages discussed above, we choose to use Mediapipe framework developed by Google which was developed for providing solutions for artificial intelligence and gesture/ moment detection problems.

IV ALGORITHM

Google's open-source media stream made its debut in June 2019. It aims to make our lives simpler by providing some integrated computer vision and machine learning technologies. One of the models in the Media Pipe Hands ML Pipeline is a palm identification model that operates across the full image and generates an oriented hand bounding box. a hand landmark model that generates very precise 3D hand key points while operating on the portion of the image that was clipped by the palm detector.

Step 1: Use camera integrated to the system and start capturing the video of the user from which the background is not processed but the user's hands are detected from the video and the rest are neglected. To perform this task, we use an opensource framework known as Media pipe.

Step 2: Media Pipe Hands uses an ML pipeline made up of several interconnected models: a model for detecting palms that uses the entire image and produces an orientated hand bounding box. a hand landmark model that produces highly accurate 3D hand key points by operating on the portion of the cropped image that the palm detector has identified.

The framework offers us the X, Y, and Z position values of the hand landmarks, which makes the data/result of tracking initially difficult to understand. Therefore, we utilize the "List" concept to list the obtained values in an understandable fashion, with each node of the list representing a value.

The corresponding landmark number, as well as its X and Y locations, are included in the list (we neglect the Z values).

Step 3: The X and Y values are distance of the hand and height at which hand is position respectively keeping camera as the origin point. We further use these positions and integrate it to the cursor so that the X and Y position is used as the input for the cursor for its next location on the screen.

Step 4: Initially there might be a lot of abrupt position changes due to fast movements of hand but this problem can be tackled by formatting the X and Y values before its given as input thereby having a gradual movement of the cursor.

Step 5: A virtual keyboard is needed to be displayed on the screen so that user's can use it instead of the traditional keyboard, and to perform the typing operation we use the same method discussed above i.e use X and Y values to hover over the keys and then perform the click operation.

V SEQUENCE DIAGRAM

A sequence diagram, commonly referred to as a system sequence diagram, is a visual representation of object interactions arranged chronologically in the discipline of software engineering. It displays the scenario's participants and the exchange of messages between them that is required for the scenario to work.

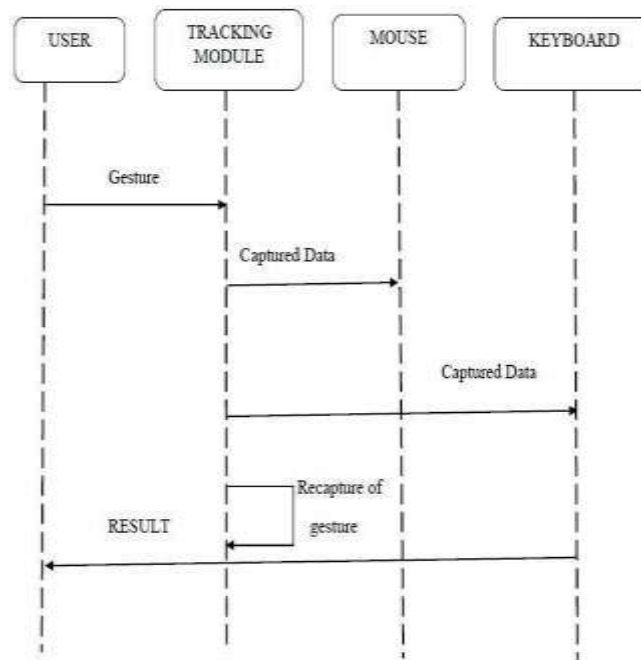


Figure 1 Sequence Diagram

VI SYSTEM ARCHITECTURE

System architecture is a conceptual representation of a system's structure, behavior, and other characteristics. [1] An architecture description is a formal description and representation of a system set up to make it easier to analyze the structures and behaviors of its constituent parts.

A system architecture may be made up of system elements and developed subsystems that will work together to accomplish the entire system. Multiple attempts have been made to formal languages that describe system architecture; these together are referred to as architectural description languages.

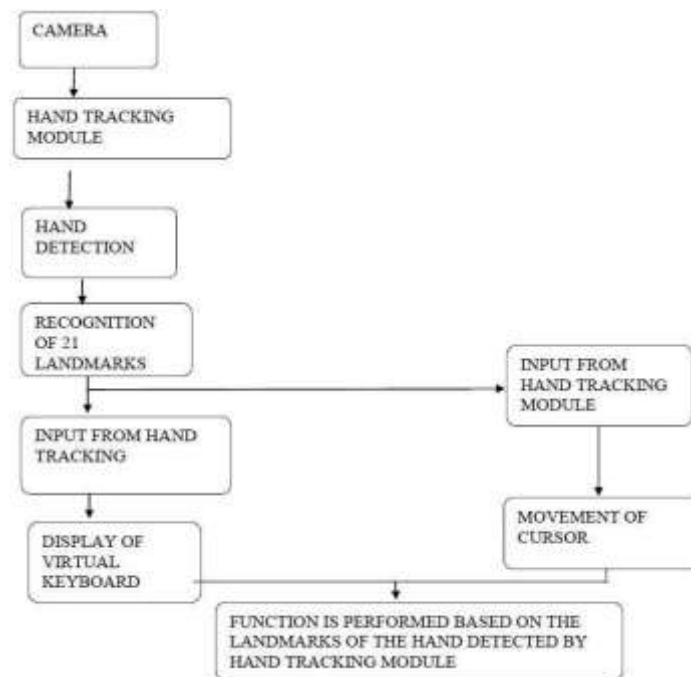


Figure 2 System Architecture

VII RESULT AND DISCUSSION

The developed project was working as intended, the response of the developed system was quick. It efficiently captured the user hand gestures by neglecting users environment and then based on user's hand or fingers location the coordinates were taken as input for the moment of mouse or click on a keyboard. Further it successfully recaptures the user's hand after recovery from human errors such as showing gestures outside the capture radius of the camera.

Further in this paper we discussed about few existing systems such as You Look Only Once (YOLO) and few of its disadvantages and also proposed a system which was best suitable for our project which is Mediapipe, a framework developed by Google.

The developed system was tested on few different laptops and computers with different configurations and the project performance was similar to that of on developed platform.



Figure 3 Hand Tracking Module

Recognizes the 21 different landmarks on palm of users using the pre-installed camera and these points will be used to detect location and helps to determine the task to be performed.

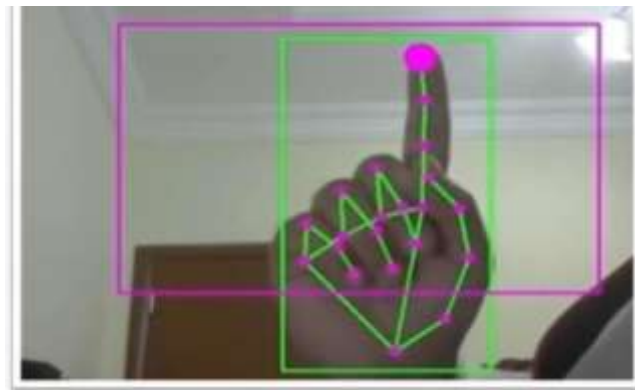


Figure 4 Virtual Mouse

Inputs the land marks from the hand recognition module and clicks on the option or icon based on these landmarks.



Figure 5 Virtual Keyboard

Inputs the land marks from hand recognition module and displays a virtual keyboard and a letter is typed based on these landmarks.

VIII CONCLUSION

The proposed Methodology which is use of Mediapipe rather than use of traditional methods or algorithms such as YOLO gave us more accurate and desired results. The developed product can further be skilled up if this method is used during the production process of laptops and desktops, if the project is installed as a software to the systems, then there would be no need of components such as mouse or keyboard. Though its not possible to

neglect the fact that both the components will still be required for users who are programmers or typewriter or in an such field because as of now the project is still in the basic stage and will not be able to fulfil the needs od such users as they require speed typing and adapting to this changed method will certainly be difficult. However, innovation of such new methods is very much essential for further progress in the field of technology and also to go hand in hand to protect environment.

IX REFERENCES

- [1] Violent behaviour detection based on SVM in elevator, G.Shu, G.Fu, P.Li And H.Geng, 2014.
- [2] Imagenet large scale visual recognition challenge, Satheesh, S.Ma, O.Russakovsky, J.Deng, H.Su, J.Krause, Z.huang, A.Khosla, A.karpathy, M.Bernstien, A.C.Berg, and L.Fei-Fie, 2014.
- [3] Vision Based Humans Detection Techniques: A Descriptive Review, Shahriar Shakir Sumit, Dayang Rohaya Amang Rambli and Seyedali Mirjalili, 2021.