International Journal of Computer Sciences and Engineering
Vol.**11**, Issue **4**, pp.**14-18**, April **2023**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

ISSN: 2347-2693 (E)

Research Paper

Hand Gesture Recognition for Human Computer Interaction through KNN Algorithm and Mediapipe

Shaik Sai Rohit[1*], Raunak Kandoi[2], Sandeep Kumar[3]

[1,2,3]Dept. of Computer Science and Engineering, Sharda School of Engineering and Technology, Sharda University, Greater Noida, Uttar Pradesh, India

**Corresponding Author: techaspire2020@gmail.com*

Received: 17/Feb/2023; **Accepted**: 20/Mar/2023; **Published**: 30/Apr/2023. | **DOI**: https://doi.org/10.26438/ijcse/v11i4.1418

Abstract - In our day to day life humans interact with computers or machines very frequently and these interactions result in completion of meaningful tasks or scheduled jobs. These interactions can involve a lot of applications like gaming, typing, scrolling, pointing, remote arm movement, etc. Out of all the ways of interacting such as mechanical movement like mouse, keyboard, joystick, etc., speech recognition, etc. The most effective one is thought to be through hand because the mechanical ones even include the movement of hand ultimately, so in order to make this interaction more convenient and efficient there was an idea of developing hand gesture recognition and it was later implemented but often it involved special instruments such as gloves with sensors or particular background. This proposed paper emphasizes the more effective way of human computer interaction which is hand gesture recognition. There are three main modules which are hand detection and hand tracking and hand gesture recognition. There are several applications of this way of interaction the user can customize to their own use. This model is fast and accurate and it can go up to 30fps and the main applications include the video game stimulation, virtual board and many other useful human computer interactions. The proposed model can detect hand even in strained backgrounds and without gloves in almost all of the cases and the model is robust and smooth. The hand gestures are one of the most natural ways of communication in humans rather than input through keyboards and mouse. This model can be used in VR and AR stimulations which would need a better way of human computer interaction than a keyboard and mouse. The main objective of this paper is to employ a new model to improve the human computer interaction. In the proposed model a menu will be displayed with the numbers representing the action desired by the user let us say 1. Represents the virtual mouse 2. Represents the virtual keyboard 3. Represents the other menu where a series of actions can be defined by the user. The proposed model uses the hand sign detection for the recognition of the numbers by the finger counting module and again using other hand sign detection module. The objective is to form a neural network to distinguish the hand signs in order recognize the hand sign to implement the desired action of the user. We have utilized Google's Mediapipe Frame Work Arrangements has further developed hand recognition model and may understand 21 3D landmarks of Palm. Subsequently we'll endeavor to know it and the method for utilizing this Library Python to understand our objective.

Keywords: Mediapipe, Sign language recognition [SLR], Human computer interaction, Gesture recognition.

1. Introduction

In the thought of growing need towards development of technology towards making the life of human easy and making the interaction between human and computer more convenient and effective there is always a need for interaction method through hands like hand gesture recognition. Since ultimately humans are more comfortable and more habituated to communicating through hands and using hands for majority of the day to day life activities. We often find hand gestures as one of the effective ways of human computer interaction. Hand gestures can be used for a variety of applications which may include a virtual board, hand communication methods, video games, human and computer interaction methods for various devices and using gesture recognitions for tasks based on user's needs. Even though there are other ways of human computer interactions such as voice recognition and image recognition, we can make a very large and complex gestures with hands and it is considered as a natural way of communication for human computer interaction. There are three main steps in this model: 1. Hand detection 2. Hand tracking 3. Gesture recognition. There are two kinds of gestures static gestures and dynamic gestures. Static gestures are recognised relatively in an easy manner while the dynamic gestures have series of gestures and the model has to recognize it through sequence of images and it has a time limit under which the gesture has to be recognised. Previously the hand gesture recognition involves wearing a hand gloves for unstrained backgrounds and proper recognition or the hand

must be in unstrained background due to its involvement in recognition of gestures, but the proposed model helps in recognizing the gestures without any unnecessary equipment and with any sort of background making it a robust and smooth model with up to 30 fps and makes it easy for hand gesture recognition and it tracks the hand and it can even be applied for virtual board using the hand as a virtual pointer. The model can have various applications in the future and can be further developed for variety of fields [10].

Sign Language Recognition attempt to foster calculations and systems for precisely unmistakable the successions of images made and understanding their which implies. A few SLR procedures misuse the matter as Gesture Recognition (GR). Consequently investigation has to this point designated on particular the positive qualities and procedures of separation in order to prop [5].

Gesture recognition could be a subject in designing additionally language innovation for the point of deciphering a singular piece with numerical algorithms. The subdiscipline of pc vision. Gestures will get back from any body development or position anyway normally appears to be on the face or hand. This concentration inside the field incorporates close to home recognition from facial and hand bit acknowledgment. Clients will utilize simple piece to direct or act with gadgets while not the piece contacting them. A few procedures are created exploitation cameras and pc vision algorithms to decipher the sign language.

Gesture based communications are dialects that utilization viewable signals to convey that implies. Communications through signing are communicated in language similarly as non-gesture based communication objects. Communications via gestures are finished regular dialects with their own synchronic etymology and vocabulary. Communications through signing don't appear to be general and don't appear to be wide perceived, however there are some draping likenesses between communications via gestures. Etymologists ponder each spoken and marked interchanges to be normal kinds of language, that implies that they each developed into a dark maturing technique, one that endured longer and advanced over the long run while not cautious planning. Language ought not be mistaken for visual correspondence, a sort of correspondence while not voice [18].

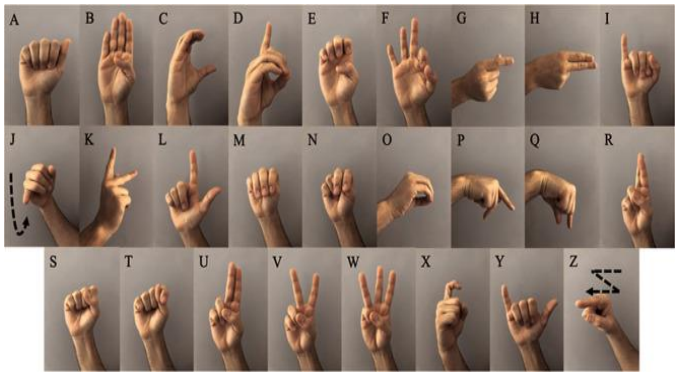

Figure1: The ASL hand sign language

MediaPipe Hands could be a solid hand and finger following gadget goal. It utilizes Machine Learning (ML) to get a handle on 21 3D local hand marks from just 1 frame. However popular ways depend for the most part on the strong work area areas for disclosure, our methodology benefits realtime execution on cell phones, even scales to a few hands. We tend to expect to permit you this convenient arrangement working on top to bottom examination and improvement society can end in instances of abuse, to push new applications and new examination ways. MediaPipe Hands utilizes partner incorporated cc pipe of the different models working together: The palm detection model that chips away at the total picture and returns the direct-coordinated hand binding box. Hand gesture model to picture cut district framed by a palm detector once returns 3D hand central issues with high liability. This system is closely resembling the one utilized in our MediaPipe Face Mesh resolution, utilizing a face detector and a face detector a milestone model .

Lately, a ton of examination has been finished on linguistic communication recognition. This acknowledgment innovation is parted into 2 classifications: - 1. Vision based Approach

The vision approach takes photographs on camera as spot data. The vision-put together methodology concentrates vigorously with respect to contact caught pictures and draws out the most and conspicuous element. Variety belts were utilized toward the beginning of the vision-based approach. The most weakness of this technique was the quality tone to be applied to the fingers. Then utilize clear hands instead of hued strips. This makes a troublesome disadvantage as these frameworks need foundation, continuous lighting, individual edges and a camera to acknowledge period execution. Moreover, such frameworks ought to be created to satisfy the necessities, as well as exactness and strength.

Hypothetical examination is predicated on anyway people comprehend information concerning their environmental elements, in any case it's very likely the first extreme to successfully utilize. Numerous totally various ways are tried up to now. The essential is to make a three-layered human hand model. The model is contrasted within reach pictures and one or extra cameras, and subsequently the boundaries tantamount to the type of the palm and consequently the consolidated points square measure measurable. These boundaries square measure then acclimated produce the piece area. The second is to require a picture abuse the camera and concentrate bound choices and individual's choices square measure utilized as contribution inside the segment algorithmic rule to isolate.

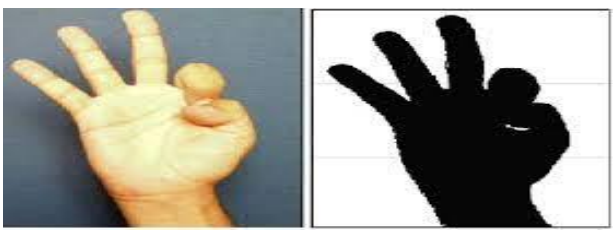
Figure2: Vision based technique sample image

The Sensor Based Approach procedure gathers information created abuse totally various sensors. The data was then examined and ends were attracted agreement with the prevalence model. Inside the instance of hand insight contrasting kinds of sensors square measure utilized and put on the hand, when the hand makes any piece, the data is recorded so broke down. The essential detecting component utilized was information gloves so LEDs. The presentation of the essential information glove was made in 1977. The sensor-based approach hinders normal hand developments as a result of the use of outside equipment. The great impediment is that exceptional contacts can't be made abuse this procedure.

In this program is that the electro mechanical component planned and controlled abuse time span digit. I the framework is made in situ of a little low regulator abuse Keil and MPLAB devices. Bit acknowledgment is performed on the objective of revision opposition is felt by the flexible detecting component. These sensors square measure associated with the hand gloves any place the information is placed in an exceptionally given framework. The framework is intended to be partitioned into 2 sections like transmitter and collector. Source the half are close by gloves any place the data sounds and handled through PIC16F7487 and sent over and over to the beneficiary class. RF innovation is utilized to move information to the recipient a piece of the two.4 GHz recurrence. ARM seven (LPC2148) processor is utilized to get information. Here from information acquired, character anticipated and matched the person to that the person is known and shown on the LCD. Contrasting sorts of courses square measure prepared of framework planned and tried continuously. Arranged the framework will be utilized for numerous applications like in vacant instrumentation, enterprises, crippled representatives and so forth.

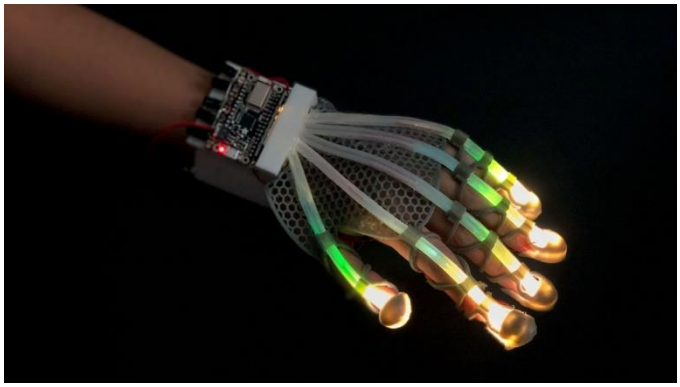

Figure3: Gloves with sensors for detection/recognition

Previously the hand gesture recognition involves wearing a hand gloves for unstrained backgrounds and proper recognition or the hand must be in unstrained background due to its involvement in recognition of gestures, but the proposed model helps in recognizing the gestures without any unnecessary equipment and with any sort of background making it a robust and smooth model with up to 30 fps and makes it easy for hand gesture recognition and it tracks the hand and it can even be applied for virtual board using the hand as a virtual pointer. The model can have various applications in the future and can be further developed for variety of fields.

2. Literature Review

Upon going through the references the advantages and disadvantages of different models are explored and the methodology of those models. Using the distance and angle as the parameters for performing the actions, if the hand is moved forward it starts the action, if moved back it stops the action, this is done by measuring the distance between camera and hand and other parameter angle is used to know if the hand is tilted towards right or left to perform certain action, the main parts are hand segmentation, hand tracking and gesture recognition [1]. The model uses the neural networks and algorithms using orientation histograms, it recognises some of the static hand gestures, it converts image into feature vector, which will be compared with the feature vectors, it is an implementation of perceptron model in MATLAB [13]. The model uses the hidden markov model and karman filter for hand gesture recognition and it can recognize 46 American sign language patterns [10]. Using of low complexity algorithms to reduce the complexity in recognition of gestures. it has three phases: 1. hand detection 2. Hand gesture recognition 3. Finger recognition and uses 2 algorithms which are : near-convex decomposition and threshold decomposition [16].

The system was built to be used for virtual keyboard and for the snake game, it uses the karman filter and compuer vision for successful hand gesture recognition, the main intention was to create a better model for human computer interaction [5]. The paper clearly discusses the trends and the different models used for hand gesture recognition which are Filtering, HMM, adaptive probabilistic model, artificial neural networks, convolutional neural networks. The paper clearly explains and compares the models with all the valid points [8].

3. Implementation Methodology

MediaPipe Hands utilizes an AI Pipeline that incorporates various collaborating models: A palm-type obtaining model that works in a really complete picture and returns a set hand-held restricting box. A penmanship model that works with a trimmed picture area framed by a palm identifier and reestablishes 3D dependable central issues.

Consequently to make a web site we must photo at least 25-30 pictures for each imprint and with this model we will get 21 hand focuses. for example joins [x, y, z]. x and y square measure normal to specify [0.0, 1.0] the aspect and level of the picture severally. The z addresses the profundity of the milestone and furthermore the profundity of the arm at the establishment, and furthermore the more modest the value the closer the camera becomes. While making the site will foresee the sign with the help of the OK Model. We'll utilize the KNN equation.
The prosed model will show a menu to the user upon running the program and the menu consist of the following:

1. Gesture 1 for virtual keyboard
2. Gesture 2 for virtual pointer
3. Gesture 3 for another menu which is user defined
4. Gesture 4 for to back to the home menu

The gestures are defined by the user for effective human computer interaction.

Hardware & Software Requirements:

1) Windows laptop or Linux, Python put in and Libraries.
2) CMOS detector (WebCam)
3) Hand bit For Visibility

Computer code we tend to accustomed acknowledge Project Signature Recognition:

1) Python put in Windows Os or Linux Os Machine.
2) Mainframe - Intel core i5 .
3) Python 3.8.6 and IDE like VS, Spyder etc.
4) Libraries: OpenCV , Tensorflow, Keras, MediaPipe and plenty of additional basic
5) KNN (The nearest neighbors) from the Sklearn Library of Python.

4. Result

This sign language recognizer can perceive signs. Every one of the signs can be perceived constant. The ongoing framework has just been prepared on a tiny data set. The proposed model showed the accurate results. The model performs the task as assumed by the user.

Figure 4a: Showing result of the proposed model

Figure 4b: Showing result of the proposed model

Figure 4c: Showing result of the proposed model

5. Conclusions and Future Scope

The proposed model with mediapipe gives a good accuracy, the average accuracy for palm detection is 95.7%. The model's accuracy is another thing to deal with, as the days pass by even the accuracy could increase and improve the model's operating nature. The model gives the results with 30FPS. More features must be added to fulfil the multiple other applications of the user which can be replaced by the human hand gestures. The future work would be improving the human computer interaction with more efficient and more accurate model.

In conclusion, MediaPipe's hand gesture detection technology holds great promise for revolutionising how we interact with computers and other gadgets. MediaPipe is perfect for a variety of applications, including gaming, virtual reality, and sign language identification, as it provides an effective and accurate method to recognise hand motions in real-time.

The fact that MediaPipe is open-source and simple to use makes it available to developers and researchers all around the world, making it a fascinating field of study with lots of room for additional investigation and advancement. The accuracy and speed of hand gesture detection utilising MediaPipe are anticipated to increase with developments in machine learning and computer vision technology, making it even more beneficial and practical in the future.

Future applications for MediaPipe-based hand gesture detection could be found in industries including healthcare, robotics, and education. For instance, employing MediaPipe for hand gesture detection in surgical settings can be used to control medical equipment without the need for the surgeon to touch anything, lowering the danger of infection and contamination. Similar to this, in the educational field, MediaPipe's hand gesture detection technology can be utilised to increase the accessibility of course materials for people with impairments.

Overall, MediaPipe-based hand gesture detection is a quickly developing technology with a lot of room for future expansion and improvement. As a result, it is a fascinating field of study that will continue to pique the curiosity of scholars, researchers, and business experts.

Conflict of Interest

There is no particular conflict of interest, the main motive was to improve the human computer interaction and improve the conditions to reduce the latency issue for hand gesture recognition and it can be applied for multiple applications in the future.

Funding Source

None

Authors' Contributions

Author-1 researched literature, conceived the study and developed the model. Author-2 involved in protocol development and gaining ethical approval. Author-3 has helped in the review and assistance of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Acknowledgements

We recognise that each and every contributor has given their consent for their recognition in this work.

REFERENCES

[1] Cristina Manresa, Javier Varona, Ramon Mas and Francisco J. Perales, "Hand Tracking and Gesture Recognition for Human-Computer Interaction" Electronic Letters on Computer Vision and Image Analysis 5(3):96-104, 2005

[2] Er. Aditi Kalsh, Dr. N.S. Garewal "Sign Language Recognition System" International Journal of Computational Engineering Research 2013

[3] Hui-Shyong Yeo & Byung-Gook Lee & Hyotaek Lim, "Hand tracking and gesture recognition system for human-computer interaction using low-cost hardware" Springer Science+Business Media New York 2013

[4] "survey on hand gesture techniques for sign language recognition."IJTRS-V2-I7-005 Volume 2 Issue VII, August 2017

[5] Kai Li1,2,3, Jun Cheng2,3, Qieshi Zhang*,2,3, Jianming Liu1, "Hand Gesture Tracking and Recognition based Human-Computer Interaction System and Its Applications" Proceeding of the IEEE International Conference on Information and Automation Wuyi Mountain, China, August 2018

[6] Bhumika Gupta, Pushkar Shukla and Ankush Mittal "K-nearest correlated neighbour classification for Indian sign language gesture recognition using feature extraction"

[7] Mais Yasen and Shaidah Jusoh, "A systematic review on hand gesture recognition techniques, challenges and applications" PeerJ computer science

[8] M S Srividya, Anala M R, "Research trends in Hand Gesture Recognition techniques" International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878 (Online), Volume-8 Issue-6, March 2020

[9] Munir Oudah , Ali Al-Naji , and Javaan Chahl, "Hand Gesture Recognition Based on Computer Vision: A Review of Techniques" Journal of Imaging July 2020

[10] Nguyen Dang Binh, Enokida Shuichi, Toshiaki Ejima, " Real-Time Hand Tracking and Gesture Recognition System" GVIP 05 Conference, 19-21 December 2005, CICC, Cairo, Egypt.

[11] Alisha Pradhan andB.B.V.L. Deepak ,"Obtaining hand gesture parameters using Image Processing",2015 International Conference on Smart Technology and Management(ICSTM).

[12] Prakash B Gaikwad, Dr. V.K.Bairagi," Hand Gesture Recognition for Dumb People using Indian Sign Language", International Journal of Advanced Research in computer Science and Software Engineering, pp:193-194, 2014.

[13] Tin Hninn Hninn Maung, "Real-Time Hand Tracking and Gesture Recognition System Using Neural Networks" World Academy of Science, Engineering and Technology 26 2009.

[14] Umang Patel & Aarti G. Ambekar "Moment based sign language recognition for IndianLanguages "2017 Third International Conference on Computing, Communication, Control and Automation (ICCUBEA).

[15] Vikram Sharma M, "Virtual Talk for deaf, mute, blind and normal humans," Texas instruments India Educator's conference, 2013.

[16] Viraj Shinde, Tushar Bacchav, Jitendra Pawar, Mangesh Sanap, "Hand Gesture Recognition System Using Camera" International Journal of Engineering Research & Technology (IJERT) Vol. 3 Issue 1, January - 2014 IJERT ISSN: 2278-0181

[17] Yellapu Madhuri and Anburajan Mariamichael)"Vision based sign language translation device" conference paper: February 2013

