International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 06 | Jun 2022 www.irjet.net p-ISSN: 2395-0072

Controlling Computer using Hand Gestures

Pradnya Kedari[1], Shubhangi Kadam[2], Rajesh Prasad[3]

[1]*Student, Department of Computer Science and Engineering, MIT School of Engineering, MIT Art, Design and Technology, Pune, Maharashtra, India*
[2]*Student, Department of Computer Science and Engineering, MIT School of Engineering, MIT Art, Design and Technology, Pune, Maharashtra, India*
[3]*Professor, Department of Computer Science and Engineering, MIT School of Engineering, MIT Art, Design and Technology, Pune, Maharashtra, India*

Abstract - *Human computer interaction platform have many ways to implement as webcams and other devices like sensors are inexpensive and can get easily available in the market. The most powerful way for communication between human and machine is through gesture. For higher conveyance between the human and machine/computer to convey information, hand gesture system is very useful. Hand gestures are a sort of nonverbal type to communicate that may be employed in several fields. Research and survey papers included hand gestures applications have acquire different alternative techniques, including those supported on sensor technology and computer vision.*

In this system, we aimed to build a real-time gesture recognition system using hand gestures. Particularly, we will use the convolutional neural network (CNN) in throughout the process. This application presents a hand gesture-based system to control a computer that is performing different operations using neural network. Our application is defined in five phases, Image frame acquisition, Hand tracking, Features extraction, Recognition of gestures and Classification (perform desired operation). An image from the webcam will be captured, and so hand detection, hand shape features extraction, and hand gesture recognition are done.

Key Words: **Deep Learning, Computer Vision, Hand Gestures, Convolutional Neural Network, Python, OpenCV**.

1. INTRODUCTION

Gesture recognition is a popular and in-demand analysis field in Human Computer Interaction technology. It has several employments in virtual environment management, medical applications, sign language translation, robot control, music creation, or home automation. There has been a special significance recently on HCI study. Hand is the one which is most helpful communication tool in several body parts, because of its expertise. The word gesture is employed for several cases involving human motion particularly of the hands, arms, and face, just some of these are informative.

The convolutional neural networks are the most popular employed technique for the image classification task. An image classifier takes an input image, or input sequence of images and categories them into one among the possible classes that it was trained to classify. They have applications in different fields such as medical domain, self-driving cars, educational domain, fraud detection, defense, etc. There are several techniques and algorithms for image classification task and also there are some challenges like data overfitting. During this project Controlling Computer using Hand Gestures, we are aimed to make a real-time application using OpenCV and Python. OpenCV is a real-time open-source computer vision and image-processing library. We'll use it with the help of the OpenCV python package.

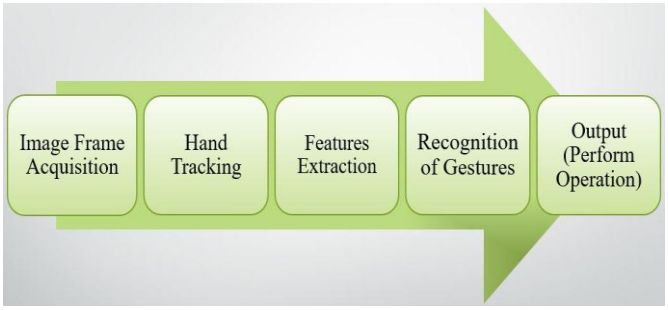

Fig 1. Methodology of Proposed System

1.1 Market Survey

Over the traditional mechanical communication technologies, gesture recognition system has become known as a most popular technology. The domain market is divided on the different basis like Technology, Type, Practice, Product, Use and Geography. Assistive robotics, Sign language detection, Immersive gaming technology, smart TV, virtual controllers, Virtual mouse, etc.

1.2 Research Gap

Most of the methods used Arduino and sensors, directly device webcam is used in very few methods. Then there might be miss-recognitions of gestures in case the background environment has elements that appears like human skin. Also hand should be within the range limit. Dataset overfitting is the main concern.

2. LITERATURE SURVEY

In computer science and technology era, gesture recognition is a crucial field which can use to translate human gesture using different computer vision techniques and algorithms. There are numerous human body motions which can create gestures but the most common type of gesture generation stand up from the face and hands. The complete policy of tracking gesture to their representation and changing them to some useful command is refer as gesture recognition. Different techniques and methods has been employed for the design and development of such kind of task.

The starting approach of interaction with computer using hand gesture was first projected by Myron W. Krueger in 1970 [1]. The aim of the perspective was achieved and also the mouse cursor control was performed with the help of an external webcam (Genius FaceCam 320), a software package that would interpret hand gestures and then turned the recognized gestures into OS commands that controlled the mouse actions on the display screen of the computer device [2]. Choosing hand gesture as a communication tool in HCI will allow the development of a good vary of applications with the absence of physical contact with the computing devices [3]. At present, most of the HCI depends on the devices such as keyboard, or mouse, however the growing importance in a category of methods and techniques based on computer vision has been popular because of skill to recognize human gestures in a simple manner [4].

Detection and recognition of a particular human body gesture and carry information to the computer is the main objective of gesture recognition. Overall objective of this system was to create the human gestures which can be admit by computer device to control a good sort of devices that are at distant using different hand poses [5]. Hand gesture recognition based on the robotic computer vision to handle the devices such as digital TV, play stations, etc. Hand gesture recognition for sign language was considered as a weighty research area lately. But because of different issues, like skin tone color difference, the complex and disturbing environments and also the different static and dynamic hand gestures, the common problem of that system raised. Hand gestures recognition for management of TV is recommended by [6]. In this, only one hand moment is used to control TV. A hand picture looks like icon that follows the hand movements of the user appear on the screen display of TV. In this paper [7], the actual human computer interface that is HCI model which is based on the hand gestures and accept gestures in a unique way to operate using monocular camera and assist the system to the HRI case has been developed. The evolved system relies on a classifier based on Convolution Neural Network to extract features and to recognize particular hand gesture.

The HMM that is Hidden Markov Model considered as a crucial tool for the dynamic gestures recognition in actual. The system employed HMM, operate in present and the general aim to build this system is to operate in static environments. The proposed methodology was for training, to employ the topology named LRB of HMM with the Baum Welch Algorithm and also for testing, the Forward and Viterbi Algorithms and checking the input sequences and building the maximum productive achievable pattern recognition state sequence [8]. In this paper [9], even the developed model seems to be easy to handle as compared to the newest available system or command based system but drawback is developed system is less powerful in spotting the gestures and recognition of the same. So despite of the complex environmental background and a normal lighting environment background, exist system need to improve and require to build further a good network for gesture recognition. This system is built for total six classes. However this exist model can be used to control operations such as power point presentation, windows picture manager, media player, games etc. In this paper [10], using an Arduino Uno and Ultrasonic sensors, operations such as handling media player, volume increase/decrease are performed on laptop. Arduino, Ultrasonic sensors, Python used for serial connection. For interactive and effective learning, such type of system can be used in the teaching classrooms.

Hand Gesture recognition system based on devices like Arduino UNO and several ultrasonic sensors to manage a device where they can control VLC by involving operations like play and pause videos and also for page scroll up and down [11]. This paper [12] suggest a convenient hand gesture monitoring system based on ultrasonic sensors, which is built using Arduino microcontroller ATMEGA32. It is claimed that extra hardware is not require to classify hand gestures and also claimed that simple low cost ultrasonic sensors can be used to notice different range limits to identify hand gestures. In this paper [13], hand gesture system relied on Arduino UNO and python programming with wired ultrasonic sensor is developed to manage a device and they included operations like zoom in/out and image rotation, etc. This trial is successful trial of working of hand motion sensing system using sensors and Arduino kits in wireless mode radio frequency. Hand gesture recognition system for Microsoft Office and media player with their own dataset is developed. [14].

3. IMPLEMENTATION AND WORKING

3.1 Dataset

We tried to use available dataset, but we faced overfitting problem. Thus we create our own dataset for training the model. We took total 10 different hand gestures to perform activities like opening the WhatsApp, PowerPoint presentation, Microsoft Word, Microsoft Edge, Google Chrome, Video Player, Xbox, and Paint etc. We took total 3000 images for training, 2000 images for testing and 500 images for validation.

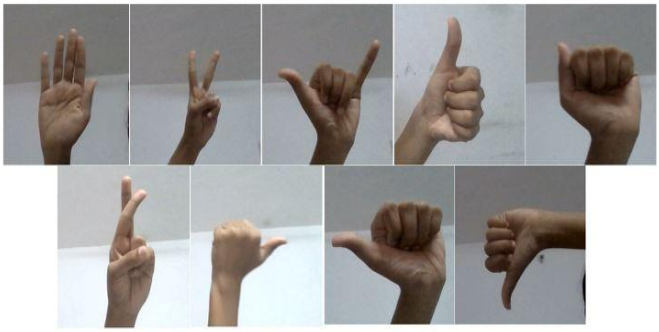

Fig 2. Sample Dataset Images

3.2 IMPLEMENTATION

a. Train The Hand Gesture Recognizer Model

In this module first upon have a glance on the dataset. We have total 3000 images for training, 2000 images for testing and 500 images for validation. We have performed image data augmentation, as somebody said "Keras ImageDataGenerator is a gem!" it lets us augment our images in real-time while the model is still training. We can apply any random transformations on each image from dataset as it is passed to the model. This will not only make our model powerful but will also lay aside on the overhead memory!

Now our next task is to train our hand gesture detector model. For that we have used Convolution Neural Network (CNN) architecture revised from Squeezenet and VGGNet. Learned features of CNN remain hidden and thus it is used as black-box.

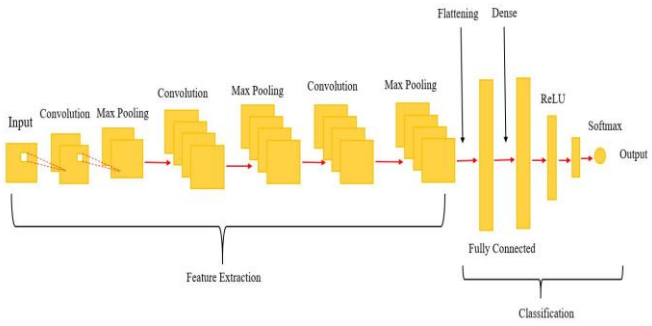

Fig 3. CNN Architecture

In this, for creating the model, we have used VGG19 (inception learning). VGG is a successor of AlexNet. VGG19 is a deep convolution neural network having convolution layer, pooling layer (max pool), fully connected layer, ReLU and softmax in its architecture. The first step is convolution operation. In this step, feature detectors are mapped, which basically serve as the neural network's filters. The second step is pooling, pooling layers are used to minimize the dimensions of the feature maps. However, will use a specific type of pooling, max pooling. Max pooling selects maximum features to the next layer. The third step will involve flattening, flatten is the function that converts the pooled feature map to a single column that is passed to the fully connected layer. Fully Connected layers, in this part, everything that we covered throughout the section will be merged together. Dense adds the fully connected layer to the neural network. In the dense layer neurons are supposed to connect deeply. Rectified Linear Unit or ReLU is a linear function that will output the input directly if it is positive, otherwise, it will output zero. Mathematically, it is defined as y = max (0, x). The softmax function transforms input value into values between 0 and 1, so that they can be interpreted as probabilities. Final output of this layer will always remain between 0 and 1. For this reason it is usual to append a softmax function as the final layer of the neural network. Softmax is often used as the final layer in the network, for a classification task.

We used Adam optimizer because accuracy rate of this optimizer is much better than others. The whole workflow of our proposed system is explained through following image Fig 4,

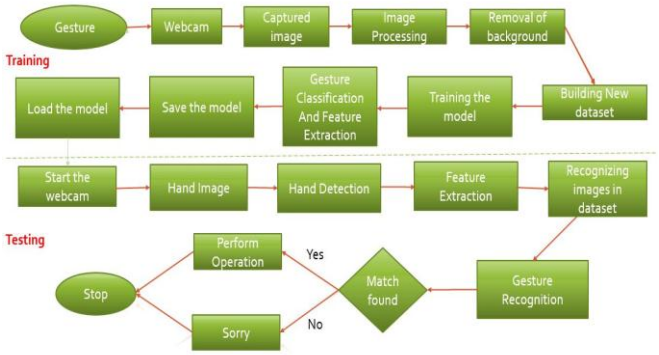

Fig 4. Workflow of Proposed System

Now after creation of model, we compiled and trained our model. Then after checking the accuracy, visualization of results by plotting graphs, and saving the model we finished with the training module. Here, complete task is taking the image of the hand gesture as an input using the webcam and then compress the image by using the CNN algorithm to match the images in the dataset in order to detect the hand moment accurately. The captured image is preprocessed, and a hand detector tries to filter out the hand image from the captured image. A CNN classifier is used to recognize gestures from the processed image after feature extraction.

b. Recognize The Hand Gesture and Perform Operations

Now we have trained model. This model will detect the hand, then features will be extracted and if gesture is recognized then particular operation will be perform. We also added one more functionality in our implementation that tells user which action is performed.

4. RESULT AND ANALYSIS

Following are the two plots of train-val accuracy and train-val loss. We got 80.40% validation accuracy.

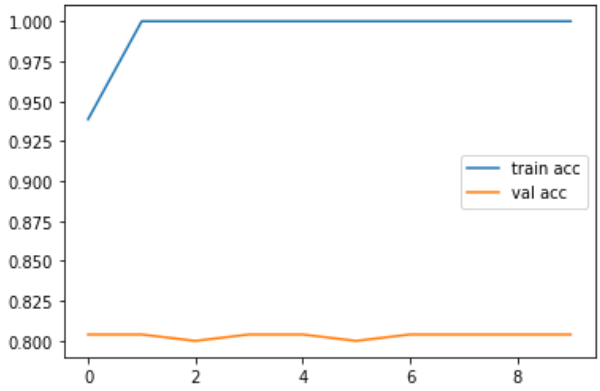

Fig 5. Train vs Val accuracy

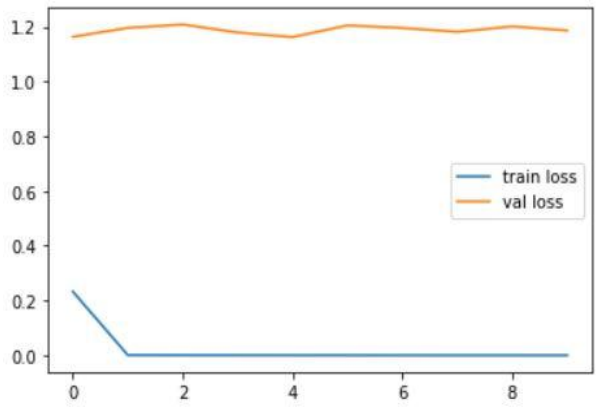

Fig 6. Train vs Val accuracy

We achieved following result of evaluation metrics. Our model is confused in some of the gestures like cross, scissor, up, etc. We achieved 85.90% accuracy at the time of testing our model. Following are the classification reports we got.

	precision	recall	f1-score	support
0	0.30	0.73	0.43	84
1	1.00	1.00	1.00	200
2	1.00	0.87	0.93	229
3	0.99	0.96	0.98	206
4	1.00	1.00	1.00	200
5	0.86	1.00	0.93	173
6	1.00	1.00	1.00	200
7	1.00	0.47	0.64	422
8	0.47	1.00	0.64	94
9	0.96	1.00	0.98	192
accuracy			0.86	2000
macro avg	0.86	0.90	0.85	2000
weighted avg	0.93	0.86	0.87	2000

Fig 7. Overall Result of Model

```
[[ 61   0   0   0   0   0   0   0  23   0]
 [  0 200   0   0   0   0   0   0   0   0]
 [  0   0 200   2   0  27   0   0   0   0]
 [  0   0   0 198   0   0   0   0   0   8]
 [  0   0   0   0 200   0   0   0   0   0]
 [  0   0   0   0   0 173   0   0   0   0]
 [  0   0   0   0   0   0 200   0   0   0]
 [139   0   0   0   0   0   0 200  83   0]
 [  0   0   0   0   0   0   0   0  94   0]
 [  0   0   0   0   0   0   0   0   0 192]]
```

Fig 8. Confusion Matrix

Below are some demons of the developed system. We assigned total 10 hand gestures to perform different operations.

Fig 9. Gesture Detecting Frame

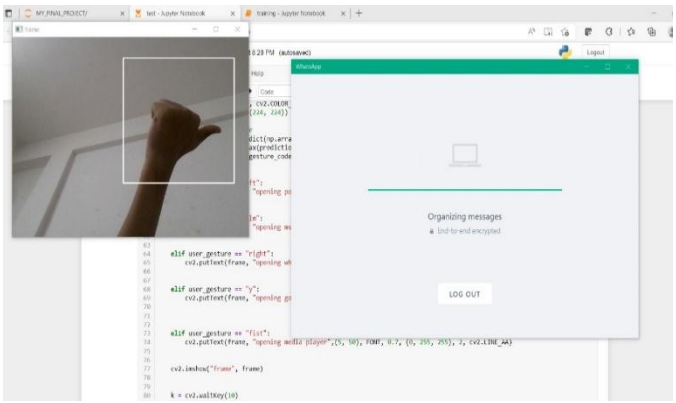

Fig 10. Opening the WhatsApp

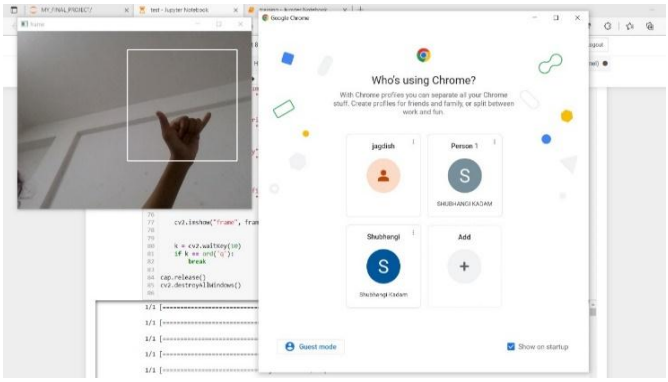

Fig 11. Opening the Google Chrome

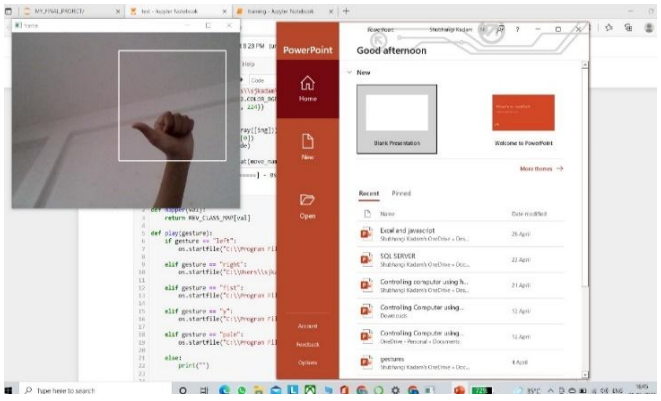

Fig 12. Opening the PowerPoint Presentation

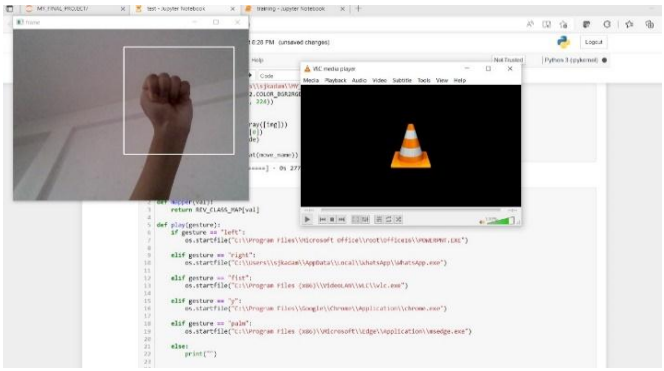

Fig 13. Opening the VLC

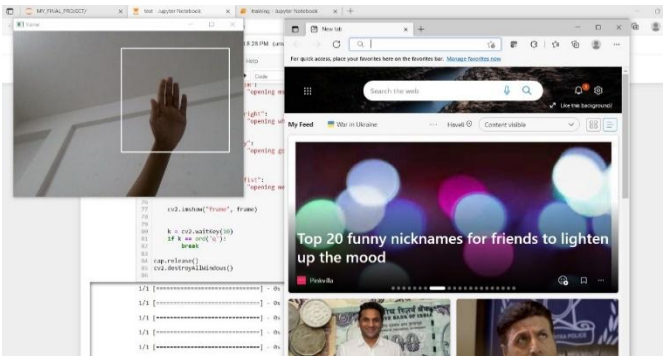

Fig 14. Opening the Microsoft Edge

5. CONCLUSION

With the growth of present technology, and as humans generally makes the use of hand movements that is hand gestures in the daily communication in order to make intentions more clear, hand gesture identification is treated to be a crucial portion of Human Computer Interaction (HCI), which provides devices the capability of detecting and classifying hand gestures, and perform activities subsequently. Research and analysis in the field of hand gestures has become more popular and exciting. It also allows a way of natural and simple interaction. Standard interactive techniques based on several tools like a mouse, keyboard/touch pad, or touch screen, joystick for gaming and consoles for system management.

In this paper we also have discussed the overall review of gesture acquisition methods, the feature extraction process, the classification of hand gestures, the challenges that face researchers in the hand gesture recognition process. In this application, we developed a deep learning model for controlling a computer using hand gestures with the help of Python and OpenCV. It is the cost-effective model as we are not using any extra devices and sensors. We can define a project as creating a suitable dataset, training the model and testing this model in real time. This project has limited scope, we assigned total 10 hand gestures to perform different operations, but in future we can add more operations like volume up/down, scroll up/down, swipe left/right and many more, and can be possible to make completely hand gestures controlling device. Hand gesture recognition used in many different applications like robotics, sign language recognition, HCI, digit and alphanumeric value, home automation, medical applications, gaming etc. Hand gestures recognition provides an interesting interaction field in a several different computer science applications.

REFERENCES

[1] M. KRUEGER, Artificial reality II, Addison-Wesley Reading (Ma), 1991.

[2] Horatiu-Stefan Grif and Cornel Cristian Farcas, "Mouse Cursor Control System Based on Hand Gesture", 9th International Conference Interdisciplinarity in Engineering, INTER-ENG 2015, 8-9 October 2015, Tirgu Mures, Romania.

[3] H.A JALAB, "Static hand Gesture recognition for human computer interaction", Asian Network for Scientific Information technology journal, pp. 1-7, 2012.

[4] J C.MANRESA, R VARONA, R. MAS and F. PERALES, "Hand tracking and gesture recognition for humancomputer interaction", Electronic letters on computer vision and image analysis, vol. 5, pp. 96-104, 2005.

[5] H. HASAN and S. ABDUL-KAREEM, "Static hand gesture recognition using neural networks", Artificial Intelligence Review, vol. 41, pp. 147-181, 2014.

[6] W. FREEMAN and C. WEISSMAN, "Television control by hand gestures", Proc. of Intl. Workshop on Automatic Face and Gesture Recognition 1995, pp. 179-183, 1995.

[7] Pei Xu, "A Real-time Hand Gesture Recognition and Human-Computer Interaction System", arXiv:1704.07296v1 [cs.CV] 24 Apr 2017.

[8] J.R Pansare, Malvika Bansal, Shivin Saxena, Devendra Desale, "Gestuelle: A System to Recognize Dynamic Hand Gestures using Hidden Markov Model to Control Windows Applications", International Journal of Computer Applications (0975 – 8887) Volume 62– No.17, January 2013.

[9] Ram Pratap Sharma and Gyanendra K. Verma, "Human Computer Interaction using Hand Gesture", Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015).

[10] Rohit Mukherjee, Pradeep Swethen, Ruman Pasha, Sachin Singh Rawat, "Hand Gestured Controlled laptop using Arduino", International Journal of Management, Technology And Engineering, October 2018.

[11] Ayushi Bhawal, Debaparna Dasgupta, Arka Ghosh, Koyena Mitra, Surajit Basak, "Arduino Based Hand Gesture Control of Computer", IJESC, Volume 10, Issue No.6, June 2020. [

[12] Udit Kumar, Sanjana Kintali, Kolla Sai Latha, Asraf Ali, N. Suresh Kumar, "Hand Gesture Controlled Laptop Using Arduino", April 2020.

[13] Sarita K., Gavale Yogesh, S. Jadhav, "HAND GESTURE DETECTION USING ARDUINO AND PYTHON FOR SCREEN CONTROL", International Journal of Engineering Applied Sciences and Technology, 2020, Vol. 5, Issue 3, ISSN No. 2455-2143, July 2020.

[14] Ram Pratap sharma, Gyanendra Varma, "Human computer interaction using hand gesture", 2015.