**Title:**

Hand Gesture Control System

**Abstract:**

Hand gesture control system is getting a great attention as it provides ability to interact with machine efficiently using human computer interaction. The main aim and objective of this research paper is to use one of the important modes of interaction that is to control the machine using hand gestures. This study recognizes the movement of hands moving in real time through a camera connected to a computer to be used for children's educational activities. The Ministry of Education of the Republic of Korea has developed a curriculum for children aged three to five through the 2019 revised Nuri curriculum so that they can be applied in the field from 2020. This project concentrates on how a system could detect, recognize and interpret the hand gesture recognition through computer vision with the challenging factors which variability in pose, orientation, location and scale. The image taken from the real-time video is analysed via Haar-cascaded Classifier to detect the gesture of hand before the image processing is done or in the other word to detect the appearance of hand in a frame. The developing of hand gesture recognition using Python and OpenCV can be implemented by applying the theories of hand segmentation and the hand detection system which use the Haar-cascade classifier. As we cannot perform so many operations using this interface as it needs a lot of CPU and GPU processing. So, for now we will understand the way and methods to achieve the target.

**Introduction:**

The main aim of building hand gesture control system is to create a natural interaction between human and computer. Where the recognized hand gesture will instruct machine to perform action instructed as per instruction. Defining human computer interaction is a term used to refer the relation between human and machine. It is also termed as Man-Machine interaction. In recent decades hand gesture recognition is considered as a new technique of Human-Computer Interaction because of its automatic, natural and easiness without requiring input from output devices like keyboard and mouse. By using Hand gestures user can interact more information in less time duration. The best advantage of hands is the ability to communicate remotely. The use of HCI hand gestures requires that the human hand configuration be measured by a computer. The main aim of proposed system is to identify specific human hand gestures and we can use it to convey information or we can control any device or robot for working applications. The purpose of this study is to help children learn how to explore nature indoors by using MediaPipe to recognize children's hand movements and use virtual mice that operate through them. In addition, in connection with the education presented in the revised Nuri curriculum, it is to find out whether there is a possibility of using MediaPipe in the remaining areas (physical exercise, health, communication, social relations) other than the field of nature exploration. The task of recognizing hand gestures is one of the main and important issues in computer vision.

The first step in any hand processing system is to detect and locate the hand in the real-time video from the webcam. The detection of hand is challenging because of variation in pose, orientation, location and scale. Also, different intensity of light in the room adds to the variability. In this study, designing of the hand gesture recognition is one of the complicated jobs that involves two major problems. Firstly, is the detection of hand. User hand is detected by using webcam in real-time video. The problem would be the unstable brightness, noise, poor resolution and contrast.

Here we are developing modules for volume control, virtual calculator, drag and drop and zoom-in and zoom-out.

**Literature Survey:**

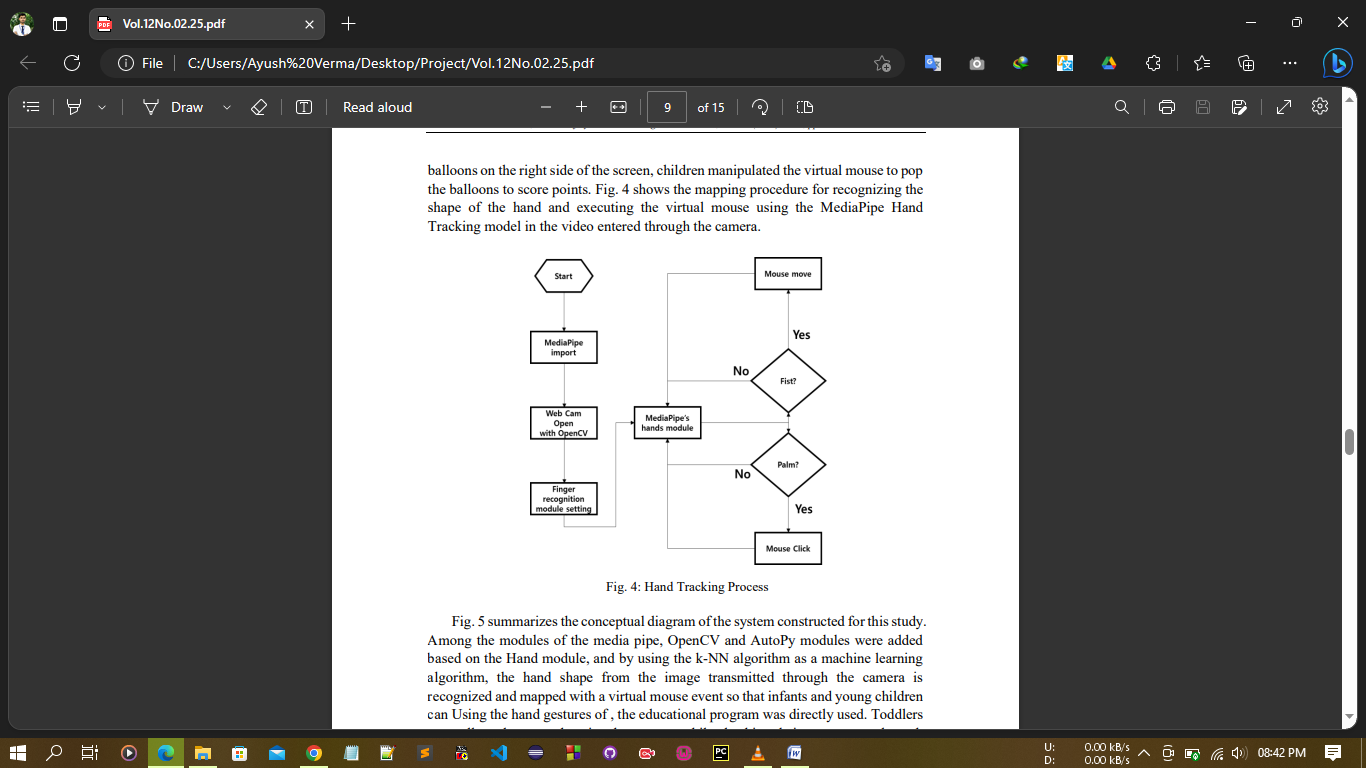
1. **1**Hasan applied multivariate Gaussian distribution to recognize hand gestures using non-geometric features. The input hand image is segmented using two different methods; skin colour-based segmentation by applying HSV colour model and clustering-based thresholding techniques. Some operations are performed to capture the shape of the hand to extract hand feature; the modified Direction Analysis Algorithm are adopted to find a relationship between statistical parameters (variance and covariance) from the data, and used to compute object (hand) slope and trend by finding the direction of the hand gesture.
2. Kulkarni recognize static posture of American Sign Language using neural networks algorithm. The input image are converted into HSV color model, resized into 80x64 and some image preprocessing operations are applied to segment the hand from a uniform background, features are extracted using histogram technique and Hough algorithm. Feed forward Neural Networks with three layers are used for gesture classification. 8 samples are used for each 26 characters in sign language, for each gesture, 5 samples are used for training and 3 samples for testing, the system achieved 92.78% recognition rate using MATLAB language.
3. P. Garg, N. Aggarwal and S. Sofa in 2009 Paper gives review about Vision based Hand Gesture Recognition methods for human computer communication, merging the different available ways, entering their common advantages and disadvantages.
4. G. Murthy and R. Jadon, Authors gives foundation of the field of gesture identification as a process for connection with computers.
5. M. K. Ahuja and A. Singh, Authors suggested a project using a database-driven hand motion identification build upon skin colour model approach and thresholding proposal further with an effectual template complement using PCA

**Research Design:**

This study was conducted according to the procedure defined in Fig. 2. Python was used as the environment for driving MediaPipe. I installed the library required for image processing in Python and the library required to use the media pipe, respectively (pip install OpenCV-python / pip install mediapipe) and configured the environment. First, the MediaPipe Hands model was applied. After importing the MediaPipe module to Python, it was confirmed that the shape and movement of the hands were normally recognized, and landmarks were applied to the finger joints. At this time, we started working after setting the configuration of MediaPipe Hand Tracking.

Second, we performed virtual hand synchronization. OpenCV was used to capture the live stream of the video transmitted from the camera installed on the laptop. OpenCV provides a very simple interface, so it is easy to use. By running the program in the Python environment, the image transmitted through the camera installed in the laptop was controlled through OpenCV and switched to mirror mode so that the recognized hand shape does not work asymmetrically.

Third, an event was applied to the virtual hand. Among the 21 landmarks provided by the MediaPipe Hands model, the program was changed to recognize only the fingers required for research. In order to perform action, a virtual volume module was mapped where the recognized hand shape was a thumb and index finger, and volume change event was mapped according to the distance between these two fingers. a virtual drag and drop module were mapped where the recognized hand shape was an index finger and middle finger, and a click event was mapped when the distance between the two fingers is too short. A virtual calculator module was designed and mapped where the recognized hand shape was a index, middle, and ring finger, and click event was mapped to select buttons, and perform operations. a virtual zoom gesture module was mapped where the recognized hand shape was a thumb and index finger, and event was mapped according to the distance between these two fingers. The thumb, index, middle, and ring fingers were recognized by landmarks 4, 8, 12, 16.

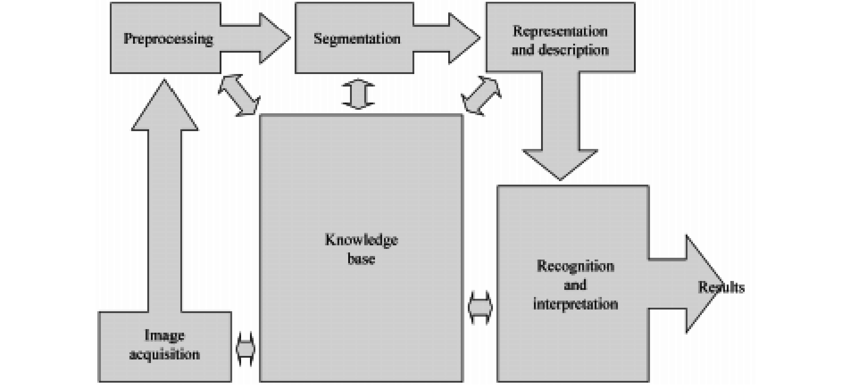


**Methodology:**

We have several steps for acquiring necessary information for performing different task. Some methods used additional hardware devices such as data gloves devices and color makers to easily extract comprehensive description of features.

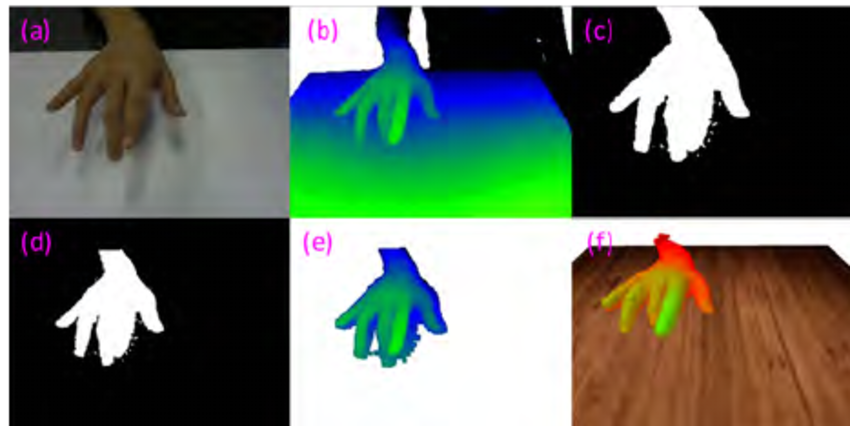
**1. Image-Processing**: Image processing(fig1) is a method to perform some operations on an image, in order to get an enhanced image and or to extract some useful information from it. If we talk about the basic definition of image processing then “Image processing is the analysis and manipulation of a digitized image, especially in order to improve its quality.”

* Importing the image
* Analysing and manipulating the image
* Output in which result can be altered image or report that is based on image analysis

****

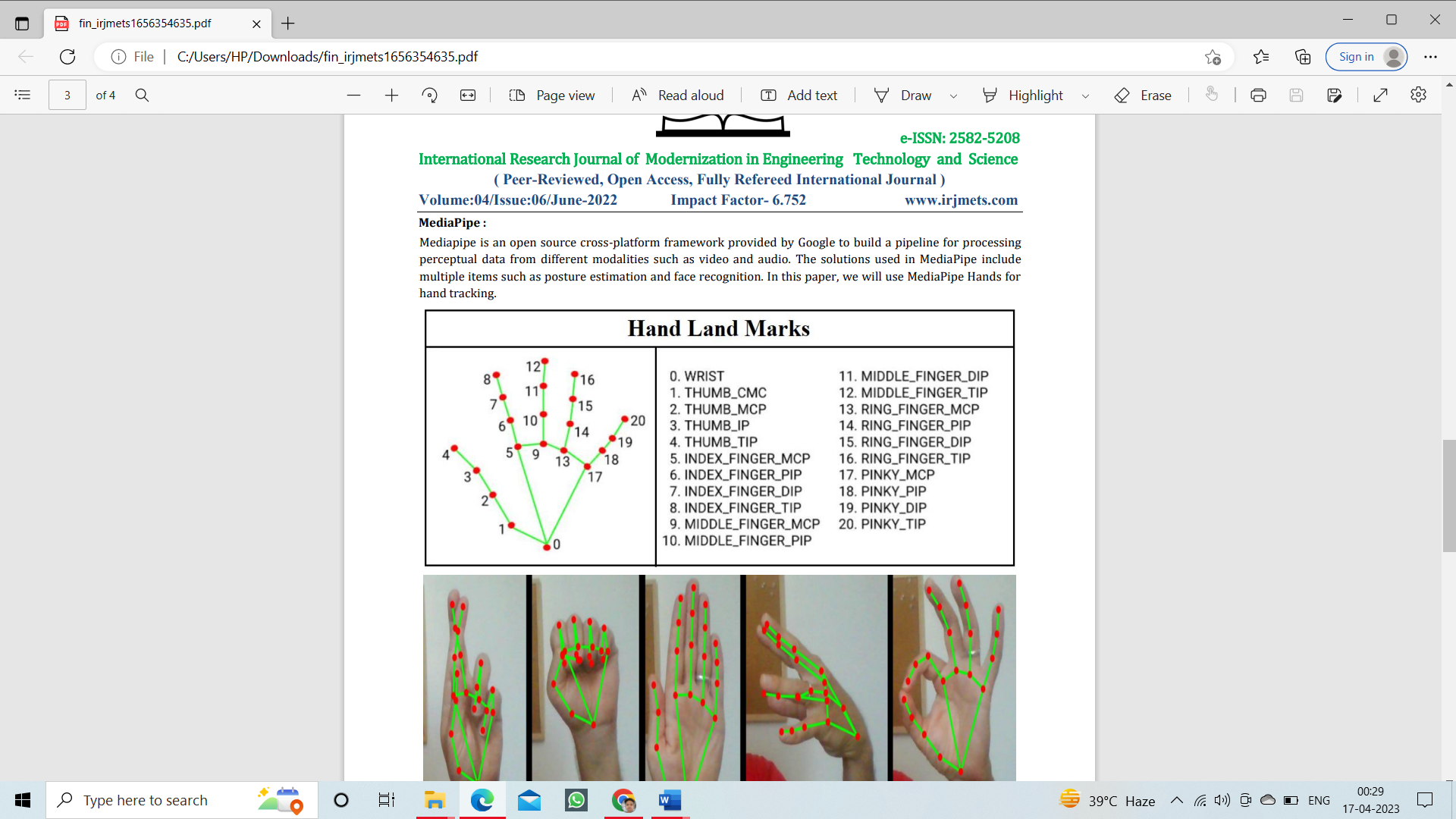
**Fig1**

**2. Segmentation**- first step for recognizing hand gestures. It is process of dividing the input image into regions separated by boundaries. If input is a dynamic gesture so we have to track and locate the hand but if it is static so we only need to segment it. As we will be using the dynamic gesture so we have to divide the video into frames and each frame has to be processed alone etc.(fig2)



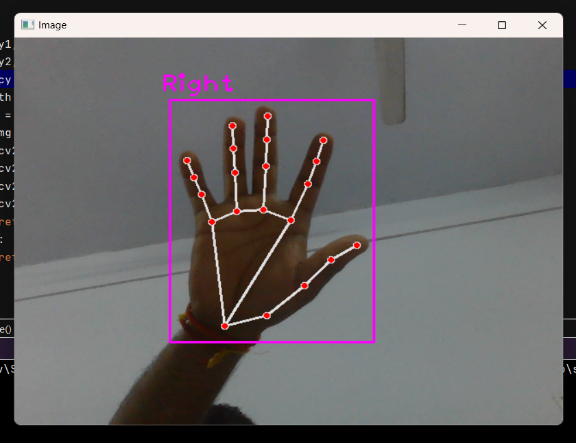
**Fig2**

**3. MediaPipe**: MediaPipe is an open-source cross-platform framework provided by Google to build a pipeline for processing perceptual data from different modalities such as video and audio. The solutions used in MediaPipe include multiple items such as posture estimation and face recognition. In this paper, we will use MediaPipe Hands for hand tracking. MediaPipe hand tracking is a model that uses a machine learning pipeline to detect and track recognized hands and fingers in images. After detecting the palm of the entire image using machine learning, as shown in Figure 3, the method uses a method of mapping 21 coordinates to define 3D coordinates to display the joints of the palm and obtain real data. MediaPipe hand tracking provides real-time performance that can be used not only on desktop computers but also on smartphones, and can be expanded to recognize multiple hands at once.



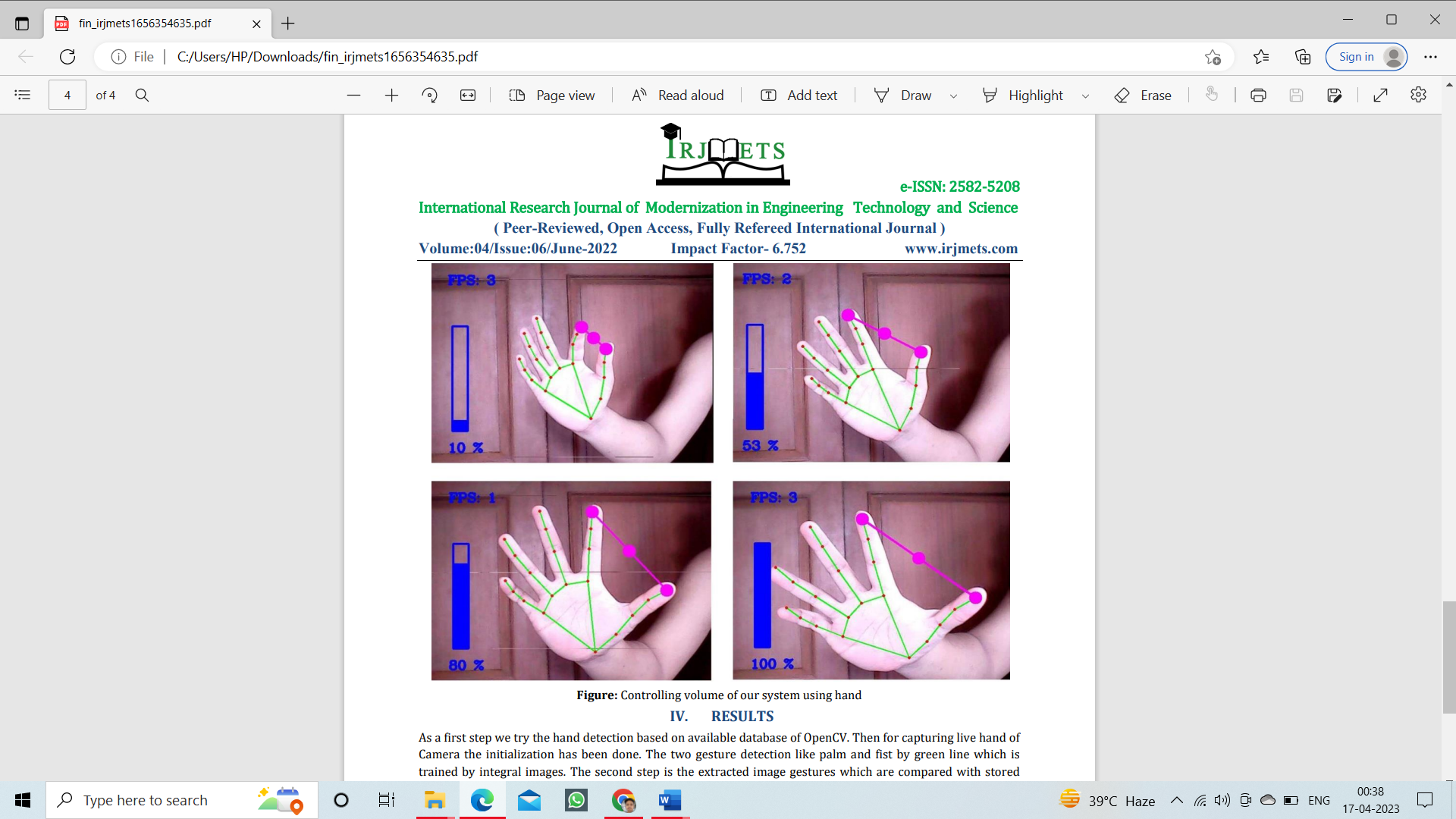
**Fig3**

**4.Feature extraction**- for extracting features we have used points on fingers and along with few points on palms that will provide us in total 21 points (fig 4). For this purpose, we can also use a google library MediaPipe used for designing pipelining for processing perceptual data from different modalities such as videos or audio.



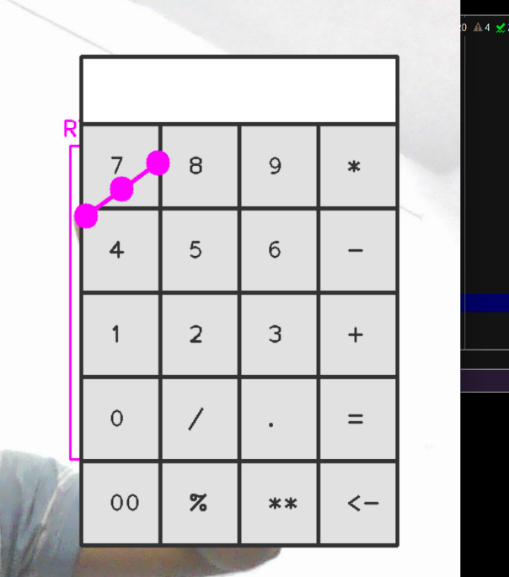
**Fig4**

**5. Pycaw:** Python Core Audio Windows Library. Pycaw is a Python library typically used in **Programming Style, Reactive Programming applications**. We can install using 'pip install Pycaw' or download it. Basically, Pycaw is used to Manipulate the Audio.(fig5)



**Fig5**

**6.Gesture classification** – After modeling and analysis of the input hand image gesture classification is used to work with for this, we have used the points of fingertips to perform the operation and the gestures for our system is difference between these fingertip points.(fig6 and fig7)

  **Fig6** **Fig7**

**Track Gesture of Hand Using Haar-Cascade Classifier:**

In order for the program to continuously detect the hand at all times, especially when the hand is in static or not moving around, the program needs to learn the hand image. This is where the machine learning comes in. In this project, the machine learning technique was performed to track the gesture of hand before the image processing is done.

**Implementation on Simulation and Hardware prototype:**

The detection of hand gesture was calculated through the space consumption within the area between the convex hull and contour of hand. Convex hull is applied to find the edge of user hand by forming descriptor in which the smallest convex set contains the edges. In this project, the OpenCV library files that is suitable for any of image processing techniques are user.

**Real-Time Object Tracking Technique:**

Due to real-time applications, modern object trackers try to reconcile as many samples as possible and keep the computation low. Kernelized correlation filter (KCF) is a variant of correlation filter. In the filter, the correlation between two samples is taken. When these samples match, the correlation value is the highest. A correlation can be found between the version of root (ROI area contains the tracked target) and ROI region at the same location in the next frame. This indicates the direction the tracked subject has moved in.

In the standard correlation filter, the following object model is not updated. If the object image changes significantly, the tracker performance decreases. In the KCF tracker, the model of the object being monitored is updated directly and continuously using the linear ridge regression model.

The process of adhering to targets using the basic KCF method consists of the following steps:

(i) Determination of grip area: it can be the initial user-defined area or an area detected by the system from the previous frame

(ii) Description of features: define the characteristics of the image area

(iii) Regression training: the detected ROI features will be added to form a dataset including past and present features as a basis for rapid training

(iv) The results after regression training are a new model, and the model is the basis for the next target detection step

The characteristics of the KCF method are relatively high accuracy, medium speed, and especially inability to recover when losing targets in a short time. Therefore, we choose the method for the paper.

**Object Classification Technique:**

Object classification techniques are particularly applied for image processing. In artificial intelligence applications, the classifier requirements are to be able to distinguish the gestures that begin with each other. The required accuracy is high and this will ensure the desired level of control accuracy. Processing speed is not required too high since information containing object is known and categorization does not occur continuously in every frame.

When the object (gesture is detected) contains the target, the next task is to recognize them. Once gesture recognition begins, we continue to track the target and end gesture. To perform the task, we choose the convolutional neural network (CNN) model. CNN is used in many problems such as image recognition, video analysis, MRI images, and natural language processing.

**Result:**

As a first step we try the hand detection based on available database of OpenCV. Then for capturing live hand of Camera the initialization has been done. The second step is the extracted image gestures which are compared with stored positive-negative integral image dataset and perform fingertip tracking by contour detection. The third step Mediapipe locate the palm and detect the 21 hand Landmarks according to Action, then we locate the points on hands which can be used for further activities.

**Conclusion:**

The Aim of this project is to develop a real time Gesture Volume Control System, real time Calculation System, real time Zoom- In and Zoom- out and real time virtual drag and drop. This paper explains a model, methods, steps, and all details needed to controlling Audio volume of system, to make a virtual calculator and its working, Zoom-In and Zoom-Out of an image and virtual drag and drop with the help of hand gestures. The proposed method here successfully developed a system which performed all these activities.

**References:**

1. [**https://www.researchgate.net/publication/284626785\_Hand\_Gesture\_Recognition\_A\_Literature\_Review**](https://www.researchgate.net/publication/284626785_Hand_Gesture_Recognition_A_Literature_Review)
2. [**https://www.researchgate.net/publication/329039460\_Research\_on\_Gesture\_Recognition\_Method\_Based\_on\_Computer\_Vision**](https://www.researchgate.net/publication/329039460_Research_on_Gesture_Recognition_Method_Based_on_Computer_Vision)
3. [**https://www.researchgate.net/publication/261339726\_Gesture\_Based\_Operating\_System\_Control**](https://www.researchgate.net/publication/261339726_Gesture_Based_Operating_System_Control)
4. [**https://www.ijtra.com/view/accessing-operating-system-using-finger-gesture-.pdf**](https://www.ijtra.com/view/accessing-operating-system-using-finger-gesture-.pdf)
5. [**https://www.irjet.net/archives/V7/i3/IRJET-V7I3926.pdf**](https://www.irjet.net/archives/V7/i3/IRJET-V7I3926.pdf)