FINAL REPORT ARTIFICIAL INTELLIGENCE

AI IN RECOGINITION (HAND GESTURE RECOGINITION)

COURSE CODE: INT404

SECTION: K21KA

TEAM MEMBER FOR THIS PROJECT

NAME	ROLL NUMBER	REGISTRATION NUMBER
Siddamsetti. Revanth Gupta	RK21KAA20	12109471
Esha	RK21KAB44	12109363
Kashinath Tiwari	RK21KAB68	12109445

SUBMITTING TO:

Ankita Wadhawan mam

TABLE OF CONTENT:

ABSTRACT:

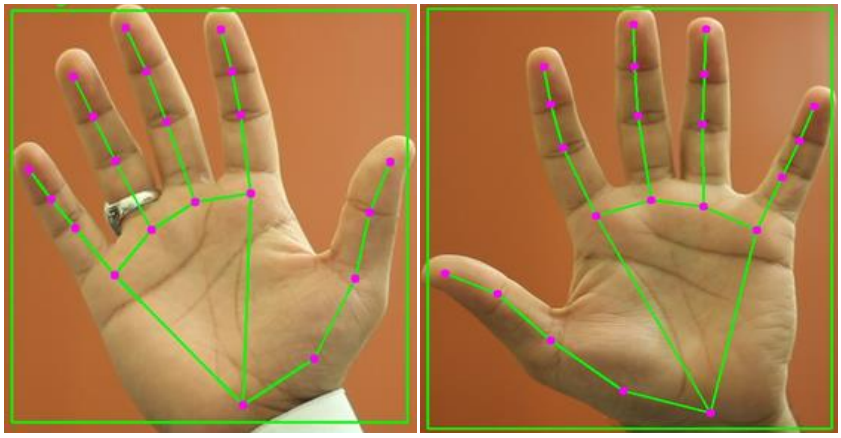
Hand gesture recognition system received great attention in the recent few years because of its manifoldness applications and the ability to interact with machine efficiently through human computer interaction. In this paper a survey of recent hand gesture recognition systems is presented.

Key issues of hand gesture recognition system are presented with challenges of gesture system. Review methods of recent postures and gestures recognition system presented as well. Summary of research results of hand gesture methods, databases, and comparison between main gesture recognition phases are also given. There are many Advantages and drawbacks are also of the Hand gesture recognition.

The project introduces an application using computer vision for Hand gesture recognition. A camera, detects hand movements in the frame, and draws lines on the image to show the movement of the hands.

Later we can trained our system for each type of count hand gestures (one, two, three, four, and five) at least once.

AI IN RECOGINITION (HAND GESTURE RECOGINITION)



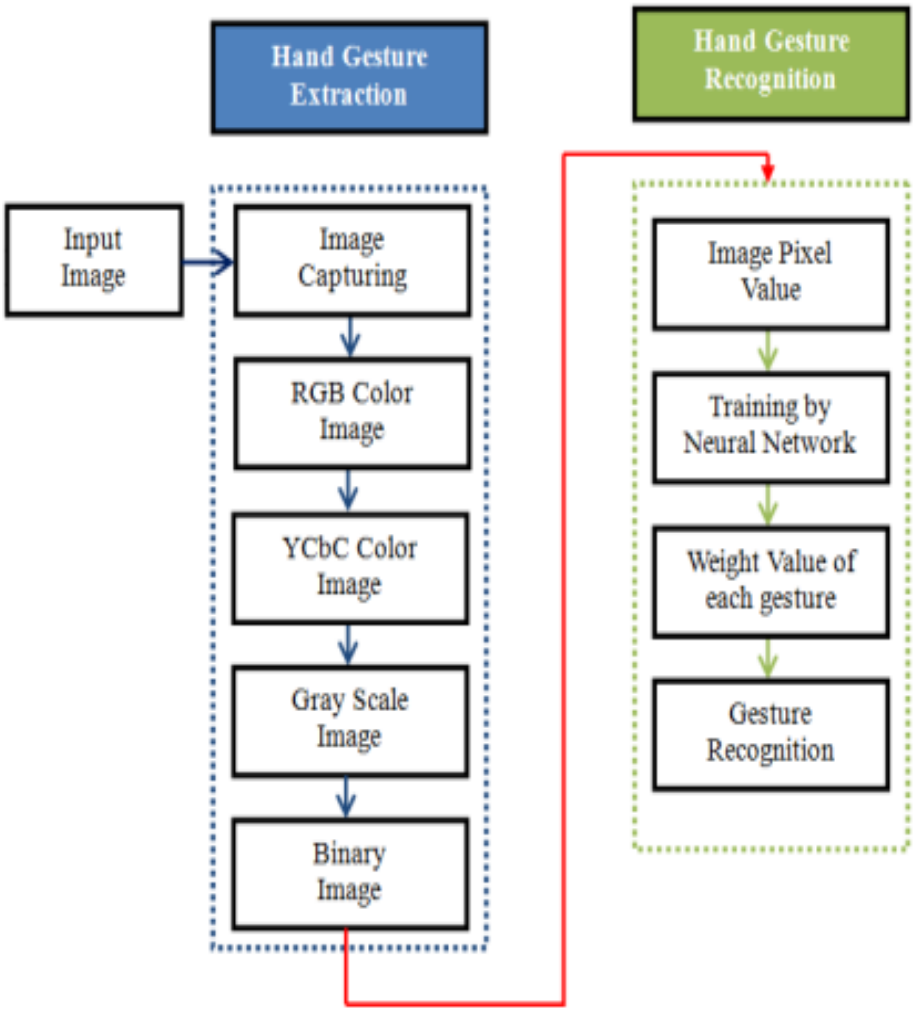
The recognition of human hand gestures by AI systems has been a valuable development over the last decade and has been adopted in high-precision surgical robots, health monitoring equipment and in gaming systems. AI gesture recognition systems that were initially visual-only have been improved upon by integrating inputs from wearable sensors, an approach known as 'data fusion'. The wearable sensors recreate the skin's sensing ability, one of which is known as somatosensory. However, gesture recognition precision is still hampered by the low quality of data arriving from wearable sensors, typically due to their bulkiness and poor contact with the user, and the effects of visually blocked objects and poor lighting. Further challenges arise from the integration of visual and sensory data as they represent mismatched datasets that must be processed separately and then merged at the end, which is inefficient and leads to slower response times. The ability to perceive the shape and motion of hands can be a vital component in improving the user experience across a variety of technological domains and platforms. For example, it can form the basis for sign language understanding and hand gesture control, and can also enable the overlay of digital content and information on top of the physical world in augmented reality. While coming naturally to people, robust real-time hand perception is a decidedly challenging computer vision task, as hands often occlude themselves or each other (e.g. finger/palm occlusions and handshakes) and lack high contrast patterns. Real-time hand gesture recognition for computer interactions is just the next step in technological evolution, and it’s ideally suited for today’s consumer landscape. Besides using gestures when you cannot conveniently touch equipment, hand tracking can be applied in augmented and virtual reality environments, sign language recognition, gaming, and other use cases.

DATASET

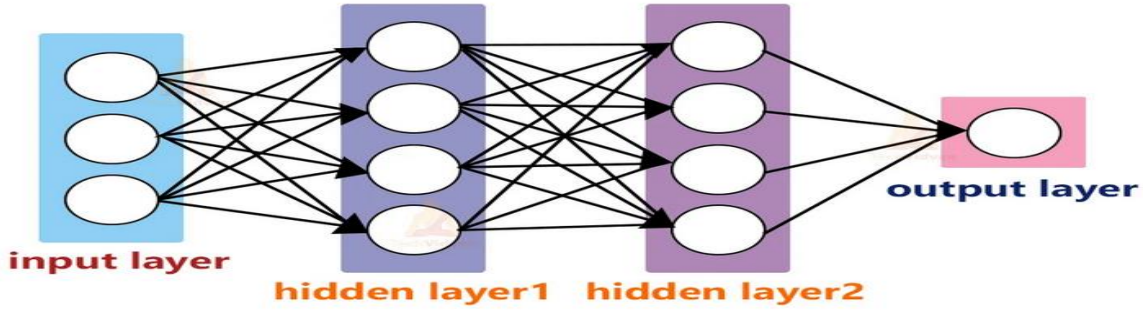
We need a dataset that contains dynamic gestures focused on interaction. There are signs where we have to use two hands and face expression. We need to synchronize them.

Link:

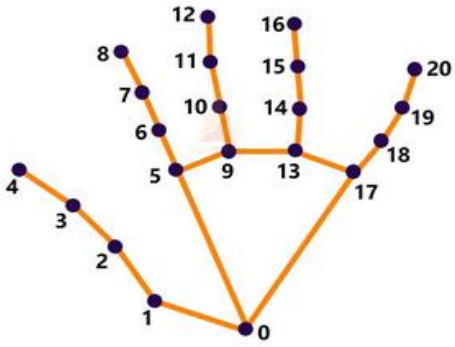
https://www.kaggle.com/datamunge/sign-language-mnist



Overall system flow



We'll first use MediaPipe to recognize the hand and the hand key points. MediaPipe returns a total of 21 key points for each detected hand.



0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

These key points will be fed into a pre-trained gesture recognizer network to recognize the hand pose.

Prerequisites for this project:

1. Python – 3.x
2. OpenCV – 4.5

3. MediaPipe – 0.8.5

4. Tensorflow – 2.5.0

5. Numpy – 1.19.3

ROLES AND RESPONSIBILITIES:

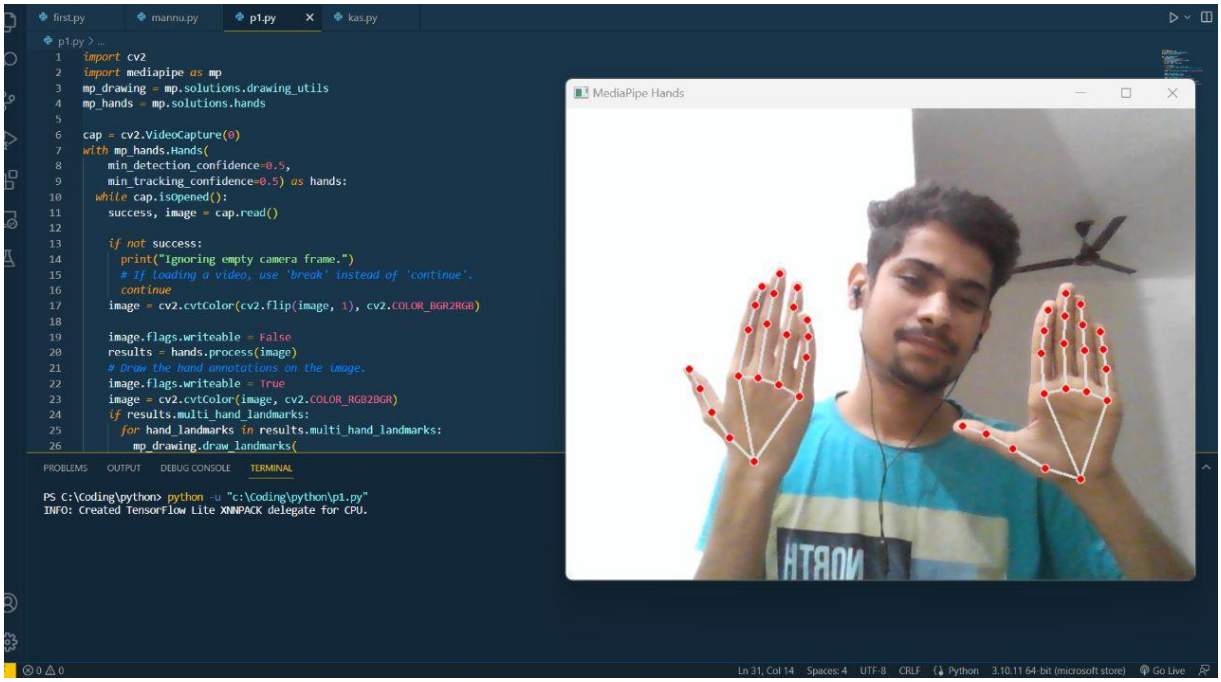
NAME	ROLES
Siddamsetti. Revanth Gupta	Project Planning, Presentation
Esha	Data Collection (using dataset), Data Analysis, PPT making
Kashinath Tiwari	Implementation of project , Report Writing

GANTT CHART

TASKS	FEBRUARY	MARCH	APRIL
Project Planning	**27**	**3**	
Data Collection and Data Analysis		**4** **15**	
Implementation of project and report writing		**16**	4

HOW CODE WORKS:

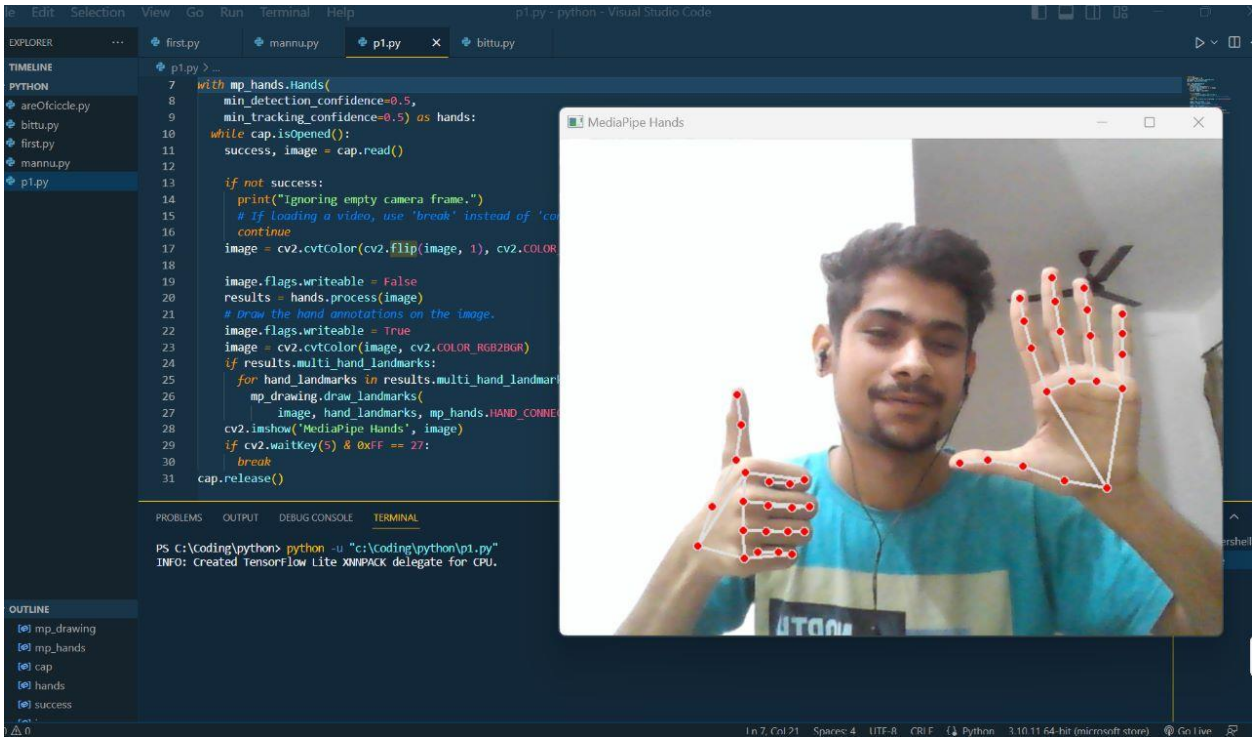
This code uses the MediaPipe Hands library and OpenCV to detect and track hand landmarks in real-time video from a camera. The first few lines import the necessary libraries: cv2 for video capturing and processing, and mediapipe for hand tracking and landmark detection. Next, cap is created as a video capture object using the default camera (device 0) as input. The with statement initializes a Hands object from mp_hands, which is used to detect and track hand landmarks in each frame. The min_detection_confidence and min_tracking_confidence parameters are set to 0.5, which means that a detection or tracking must have at least a 50% confidence level to be considered valid. The main loop begins with while cap.isOpened():, which means that the loop will continue as long as the video capture object is successfully opened. Inside the loop, success, image = cap.read() reads a frame from the video capture object and assigns it to image. If success is False, the loop continues to the next iteration. image is then flipped horizontally using cv2.flip() and converted from BGR format to RGB format using cv2.cvtColor(). image.flags.writeable = False prevents the image from being modified in-place, and results = hands.process(image) detects and tracks hand landmarks in the image using the Hands object from mp_hands. After the landmarks have been detected and tracked, image.flags.writeable = True allows the image to be modified in-place, and image = cv2.cvtColor(image, cv2.COLOR_RGB2BGR) converts the image back to BGR format. Finally, if results.multi_hand_landmarks is not None, the hand landmarks and connections are drawn on the image using mp_drawing.draw_landmarks(). The resulting image is displayed in a window named 'MediaPipe Hands' using cv2.imshow(). The loop continues until the 'Esc' key is pressed, which is detected by if cv2.waitKey(5) & 0xFF == 27:. Finally, cap.release() releases the video capture object and frees any resources used by it.

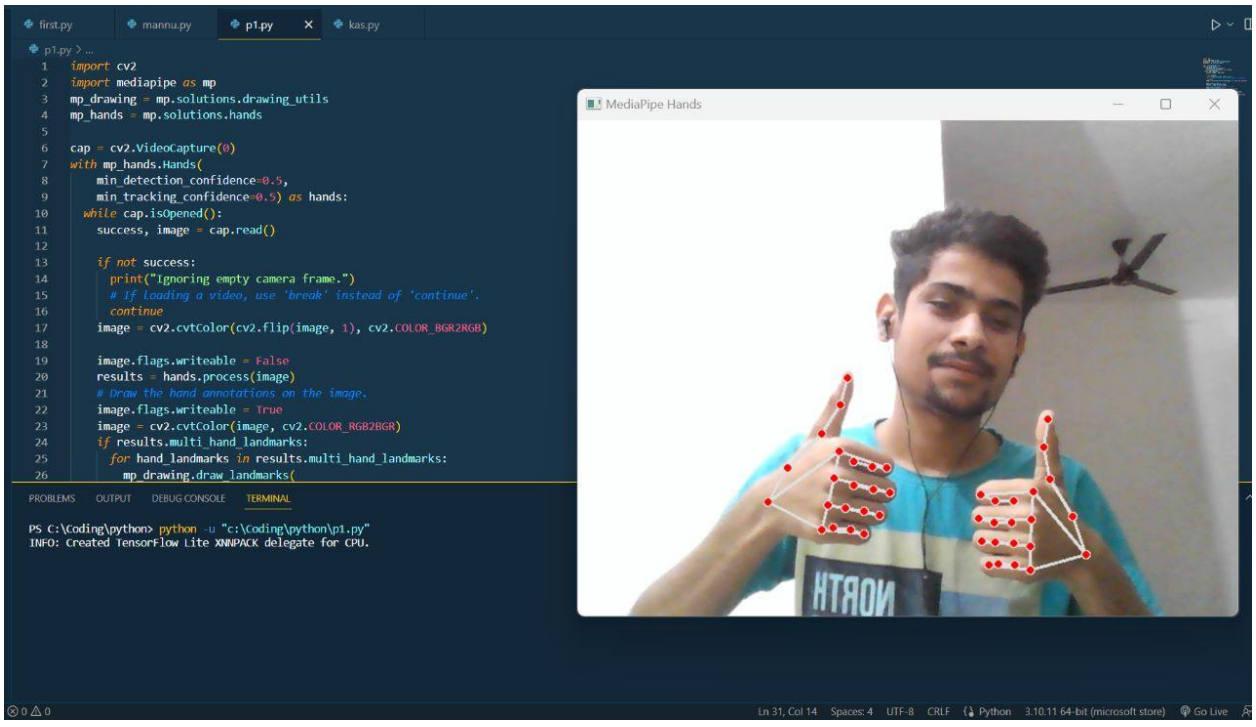


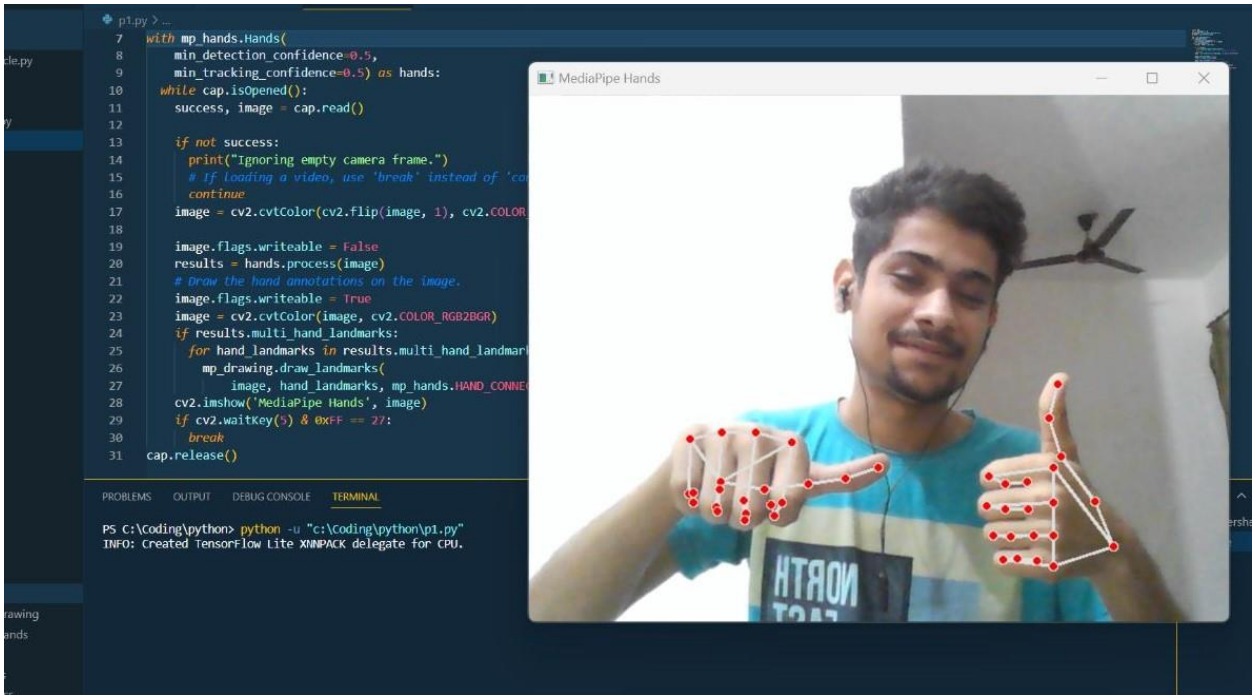
IMPLEMENTATION OF PROJECT:

We have done the project Hand Gesture Recognition using Python. In the Python code we have used a library called MediaPipe Hands and OpenCV to detect and track hand movements in real-time video from the computer's camera. The Python code in the project first sets up a connection to the computer's camera and initializes the hand tracking system using the MediaPipe Hands library. It then starts a loop that reads each frame of video from the camera, detects hand movements in the frame, and draws lines on the image to show the movement of the hands. The loop continues until the user presses the 'Esc' key, at which point the program stops tracking hand movements and exits. Overall, the code allows you to visualize how your hands move in real-time video, which can be helpful for studying hand movements or creating interactive programs that respond to hand gestures.

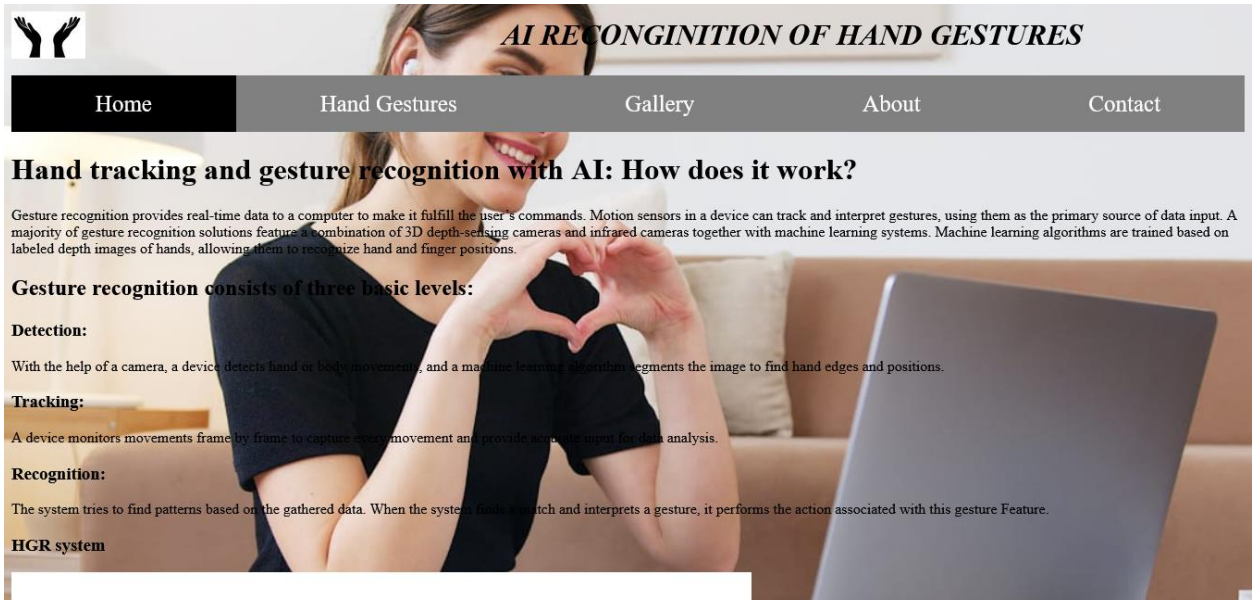
IMAGES OF WORKING OF HAND GESTURE RECOGNITION:





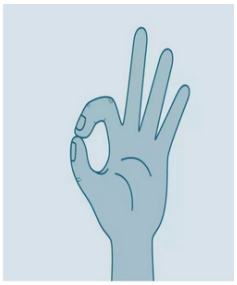


IMAGES OF WORKING WEBSITE OF HAND GESTURE:



HOW HAND GESTURES ARE USED IN OUR DAILY LIFE

Okay Symbol



A-OK or Okay, made by connecting the thumb and forefinger in a circle and holding the other fingers straight, usually signal the word okay.Gestures are a form of nonverbal communication in which visible bodily actions are used to communicate important messages, either in place of speech or together and in parallel with spoken words.Gestures include movement of the hands, face, or other parts of the body.

Peace Symbol

The 'V' sign, made by holding up the index and middle fingers, initially was used to signal victory by Allied nations during World War II. Anti-war activists later adopted it as a symbol of peace, and today the gesture is known as "the peace sign."

Thumbs up is perhaps the most common of hand gestures and one that has been used for thousands of years. The signal for approval or agreement is commonly found European and American cultures as a sign of approval or that things are going according to plan. The gesture is so prevalent that it is a common emoji and is commonly used in social media and customer service ratings to indicate satisfaction.

However, in many Islamic and Asian countries, it is considered a major insult. In Australia, the gesture also means all is fine unless the user moves it up and down which transfers the gesture to an insult.

The thumbs down gesture is also commonly used in America, but less so in many other countries. The gesture obviously means the opposite of “thumbs up”, however in many cultures it is considered to be very rude and arrogant.

Thumbs up is perhaps the most common of hand gestures and one that has been used for thousands of years. The signal for approval or agreement is commonly found European and American cultures as a sign of approval or that things are going according to plan.

A-OK or Okay, made by connecting the thumb and forefinger in a circle and holding the other fingers straight, usually signal the word okay.Gestures are a form of nonverbal communication in which visible bodily actions are used to communicate important messages, either in place of speech or together and in parallel with spoken words.Gestures include movement of the hands, face, or other parts of the body.

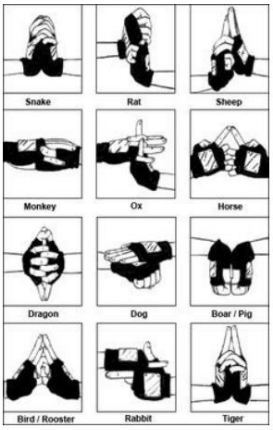
The 'V' sign, made by holding up the index and middle fingers, initially was used to signal victory by Allied nations during World War II. Anti-war activists later adopted it as a symbol of peace, and today the gesture is known as "the peace sign."

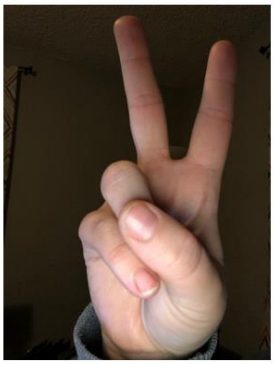
Two handed signs

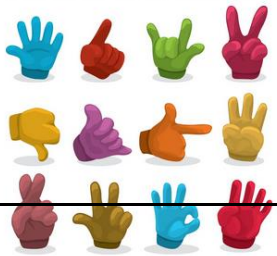
Air quotes are made by raising both hands to eye level and flexing the index and middle fingers of both hands while speaking. Their meaning is similar to that of scare quotes in writing. Añjali Mudrā (namaste) is a sign of respect in India and among yoga practitioners. It is made by pressing the palms together. This is of Indian origin. Applause is an expression of approval made by clapping the hands together to create repetitive staccato noise. Applause is most appropriate within a group setting, to collectively show approval by the volume, duration, and clamor of the noise. Awkward turtle is a two handed gesture used to mark a moment as awkward. One hand is placed flat atop the other with both palms facing down, fingers extended outward from the hand and thumbs stuck out to the sides. The thumbs are rotated to symbolize flippers.[35] Batsu. In Japanese culture, the batsu (literally: ×-mark) is a gesture made by crossing one's arms in the shape of an "X" in front of them in order to indicate that something is "wrong" or "no good".[36] Bras d'honneur is an obscene gesture made by flexing one elbow while gripping the inside of the bent arm with the opposite hand. Arms crossed on chest - denotes suspicion or guardedness.

The Kohanic or Priestly Blessing – a gesture of benediction in Judaism, used (especially by those of Kohanic or priestly descent) when reciting the Priestly Blessing (Numbers 6: 22-26). Both hands are held up, palms toward the congregation, with the fingers grouped in twos – the little and ring fingers together, the index and second fingers together, and the tips of the two thumbs touching. The golf clap, unlike applause, is a timid and

Images on Hand Gestures





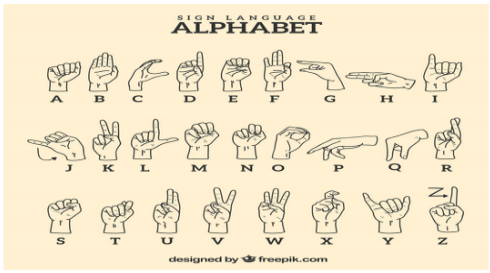


what is Hand Gestures

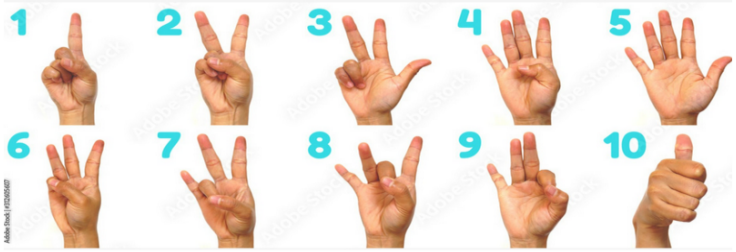
Gestures are a form of nonverbal communication in which visible bodily actions are used to communicate important messages, either in place of speech or together and in parallel with spoken words.Gestures include movement of the hands, face, or other parts of the body. Physical non-verbal communication such as purely expressive displays, proxemics, or displays of joint attention differ from gestures, which communicate specific messages.Gestures are culture-specific and may convey very different meanings in different social or cultural settings.Hand gestures used in the context of public speaking or musical conducting are chironomy,while those used in an elaborate system of specific communication are known as sign language. Although some gestures, such as the ubiquitous act of pointing, differ little from one place to another, most gestures do not have invariable or universal meanings, but connote specific meanings in particular cultures. A single emblematic gesture may have very different significance in different cultural contexts, ranging from complimentary to highly offensive. This list includes links to pages that discuss particular gestures, as well as short descriptions of some gestures that do not have their own page. Not included are the specialized gestures, calls, and signals used by referees and umpires in various organized sports. Policemen also make gestures when directing traffic. Miming is an art form in which the performer uses gestures to convey a story; charades is a game of gestures. Mimed gestures might generally be used to refer to an action in context, for example turning a pretend crank to ask someone to lower a car side window (or for modern power windows, pointing down or miming pressing a button).

Single handed

Fingerspelling in sign language



Fingerspelling has been introduced into certain sign languages by educators and as such has some structural properties that are unlike the visually motivated and multi-layered signs that are typical in deaf sign languages. In many ways fingerspelling serves as a bridge between the sign language and the oral language that surrounds it. Fingerspelling is used in different sign languages and registers for different purposes. It may be used to represent words from an oral language that have no sign equivalent or for emphasis or clarification or when teaching or learning a sign language.
[1]In American Sign Language (ASL) more lexical items are fingerspelled in casual conversation than in formal or narrative signing.
[2] Different sign language speech communities use fingerspelling to a greater or lesser degree. At the high end of the scale
[3] fingerspelling makes up about 8.7% of casual signing in ASL[2] and 10% of casual signing in Auslan.
[4] The proportion is higher in older signers. Across the Tasman Sea only 2.5% of the corpus of New Zealand Sign Language was found to be fingerspelling.
[5] Fingerspelling did not become a part of NZSL until the 1980s.
[6] Before that words could be spelled or initialised by tracing letters in the air.
[7] Fingerspelling does not seem to be used much in the sign languages of Eastern Europe except in schools.
[8] and Italian Sign Language is also said to use very little fingerspelling, and mainly for foreign words. Sign languages that make no use of fingerspelling at all include Kata Kolok and Ban Khor Sign Language.
The speed and clarity of fingerspelling also vary among different signing communities. In Italian Sign Language fingerspelled words are produced relatively slowly and clearly, whereas fingerspelling in standard British Sign Language (BSL) is often rapid so that the individual letters become difficult to distinguish and the word is grasped from the overall hand movement. Most of the letters of the BSL alphabet are produced with two hands but when one hand is occupied the dominant hand may fingerspell onto an imaginary subordinate hand and the word can be recognised by the movement. As with written words, the first and last letters and the length of the word are the most significant factors for recognition. When people fluent in sign language read fingerspelling they do not usually look at the signer's hand(s) but maintain eye contact, as is normal for sign language. People who are learning fingerspelling often find it impossible to understand it using just their peripheral vision and must look straight at the hand of someone who is fingerspelling. Often they must also ask the signer to fingerspell slowly. It frequently takes years of expressive and receptive practice to become skilled with fingerspelling.



Chinese number gestures are a method to signify the natural numbers one through ten using one hand. This method may have been developed to bridge the many varieties of Chinese—for example, the numbers 4 (Chinese: 四; pinyin: sì) and 10 (Chinese: 十; pinyin: shí) are hard to distinguish in some dialects. Some suggest that it was also used by business people during bargaining (i.e., to convey a bid by feeling the hand gesture in a sleeve) when they wish for more privacy in a public place. These gestures are fully integrated into Chinese Sign Language.

MENTODS

While the five digits on one hand can easily express the numbers one through five, six through ten have special signs that can be used in commerce or day-to-day communication. The gestures are rough representations of the Chinese numeral characters they represent. The system varies in practice, especially for the representation of "7" to "10". Two of the systems are listed below:

Zero (〇)

Northern China: The fist is closed. This may be interpreted as 10 depending on the situation, though some Chinese distinguish between zero and ten by having the thumb closed or open, respectively. Coastal southern China: The thumb and index finger make a circle, with the other three fingers closed.

One (一)

The index finger is extended.

Two (二)

The index and middle fingers are extended.

Three (三)

The thumb and index finger are closed and the other three fingers are extended. The thumb holds the little finger down in the palm and the middle three fingers are extended.

Four (四)

The thumb is held in the palm and the four fingers are extended.

Five (五)

All five digits are extended. Only the thumb is extended (either upwards or outwards) with the palm facing the signer. Counting with fingers is often different from expressing a specific number with a finger gesture. When counting, the palm can be either facing its owner or the audience, depending on the purpose. Before counting, all fingers are closed; counting starts by extending the thumb as the first, then the index finger as the second, until all fingers are extended as the fifth; then counting can be continued by folding fingers with the same sequence, from thumb through the little finger, for counting from the sixth through the tenth. Repeating the same method for counting larger numbers. One can also starts counting with all fingers extended. Some believe that for formal scenario such as giving speech or presentation, counting with the palm facing the audience and starting with all fingers extended is more polite, since the gesture of folding of fingers representing bowing. When playing drinking finger games (划拳, 猜拳), slightly different sets of finger gestures of numbers is used. One of them is:

Zero (〇)

The fist is closed.

One (一)

The thumb is extended with all other fingers folded toward the palm.

Two (二)

The thumb and index finger make an "L", other fingers closed.

Three (三)

With the last two fingers closed and the rest fingers (the thumb and the first two fingers) extended, or With the index finger and thumb closed, the last three fingers are extended.

Four (四)

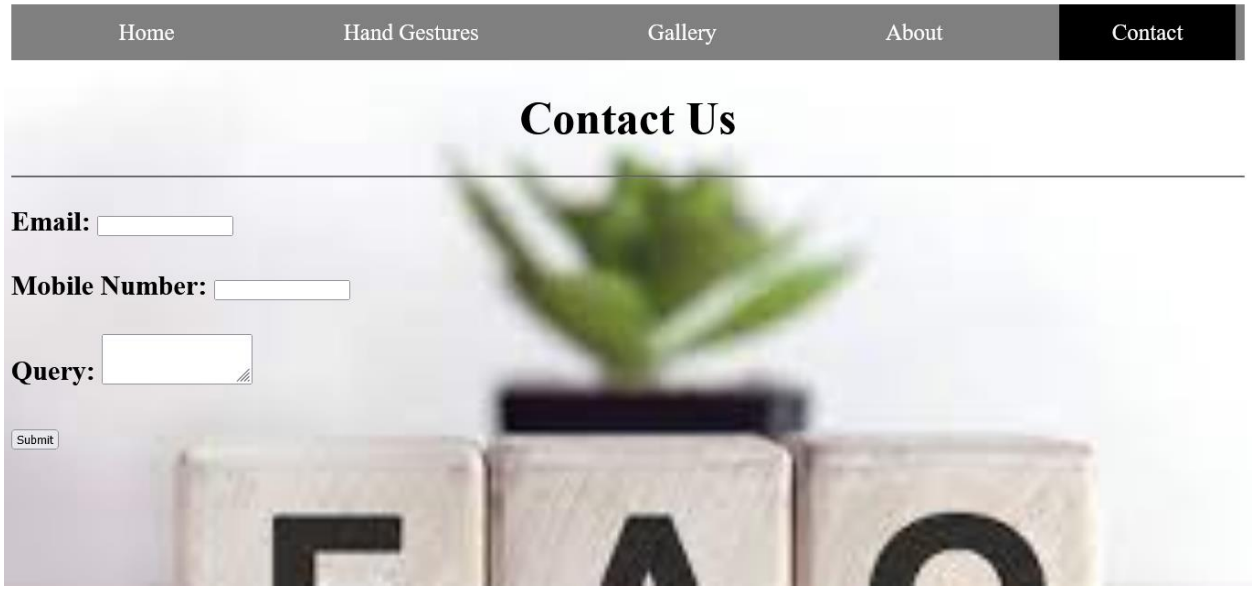
The thumb is held in palm with the four fingers extended.

Five (五)

All five digits are extended.

Six (六)

The little finger and thumb are extended (the extended thumb indicating one set of 5); the other fingers are closed, sometimes with the palm facing the



CONCLUSION:

This code is an example of how to use the MediaPipe Hands library for hand tracking and recognition in Python. First, the necessary libraries are imported: OpenCV for capturing video from the camera and displaying images, and MediaPipe Hands for hand tracking and recognition. Next, the camera is opened with cv2.VideoCapture(0), which captures video from the default camera. Inside the while loop, each frame of the video is read using cap.read(), and then converted to RGB format and passed to the MediaPipe Hands library for processing. The hands.process(image) method detects the hand landmarks in the image and returns the results. If any hand landmarks are detected, they are drawn on the image using the mp_drawing.draw_landmarks() method, which takes the image, hand landmarks, and hand connections as input. Finally, the processed image is displayed using cv2.imshow(), and the program exits when the Esc key is pressed. This code can be extended to perform hand gesture recognition by training a machine learning model on a dataset of labeled hand gesture images, and using the model to classify the detected hand gestures. Additionally, the hand landmarks detected by MediaPipe Hands can be used for various hand-based computer vision applications, such as hand gesture control of a computer interface or virtual reality environment.

REFERENCE:

https://www.geeksforgeeks.org/project-idea-dynamic-hand-gesture-recognition-using-neural-network/

https://www.javatpoint.com/hand-gesture-controlled-robot

https://www.researchgate.net/publication/284626785_Hand_Gesture_Recognition_A_Literature_Review