



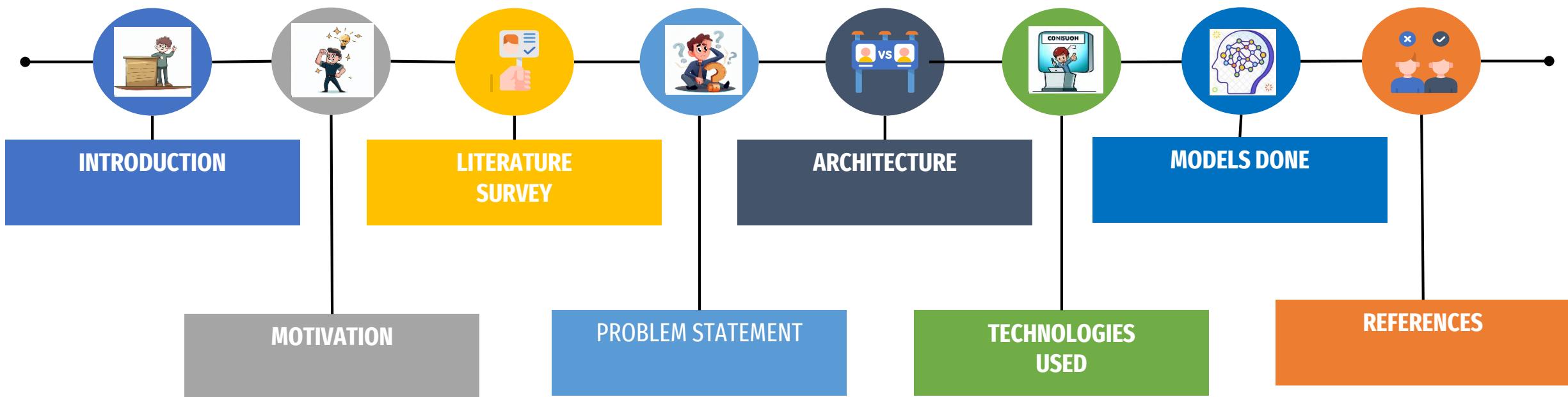
# REALTIME SIGN LANGUAGE TO VOICE

## USING DEEP LEARNING

**GUIDED BY**  
**DR. VICTER**  
**PAUL**

**STUDENT**  
**KAVALI SOFIA**  
**SAGAR**

# CONTENTS



# Introduction

Thinking is directly correlated to language. Language provides the framework for an individual's thoughts, as well as for society, values, and beliefs.

Reading presents a significant challenge for individuals who are born deaf because they cannot hear the language that is encoded by print.

Because written language learning is typically linked to hearing and producing language, a student who is deaf has an added barrier to learning any written language

Of four language skills: Listening, Speaking, Reading and Writing, first two are related to sound and later two to light.

Blind people learn Braille writing system for reading and writing to understand and adjust in the world.

Deaf and Dumb people use sign language to communicate which has some inherent issues to learn about the world.

# Motivation



Deaf and Dumb people can see the entire world but they cannot communicate with others by languages which were invented by humans.

They use sign language which is natural to posses by observing the actions and movements happen in the surrounding world.

At present more than 300 Sign languages are there in the world. People who knows Sign language can communicate with each other. But the communication-gap between the deaf and normal people persists.

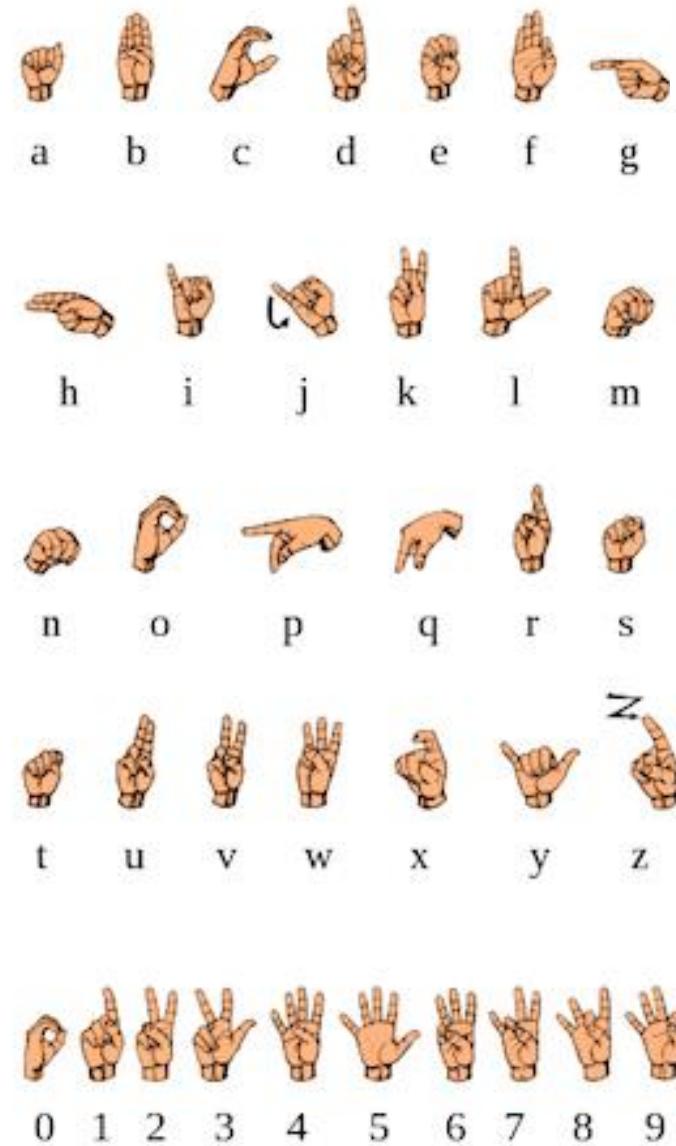
Born deaf people do not know how to learn that normal children acquire by sound-instructions and this leads to language learning-impairment at latter stages and makes them depend on Sign language throughout their life

# Literature Survey

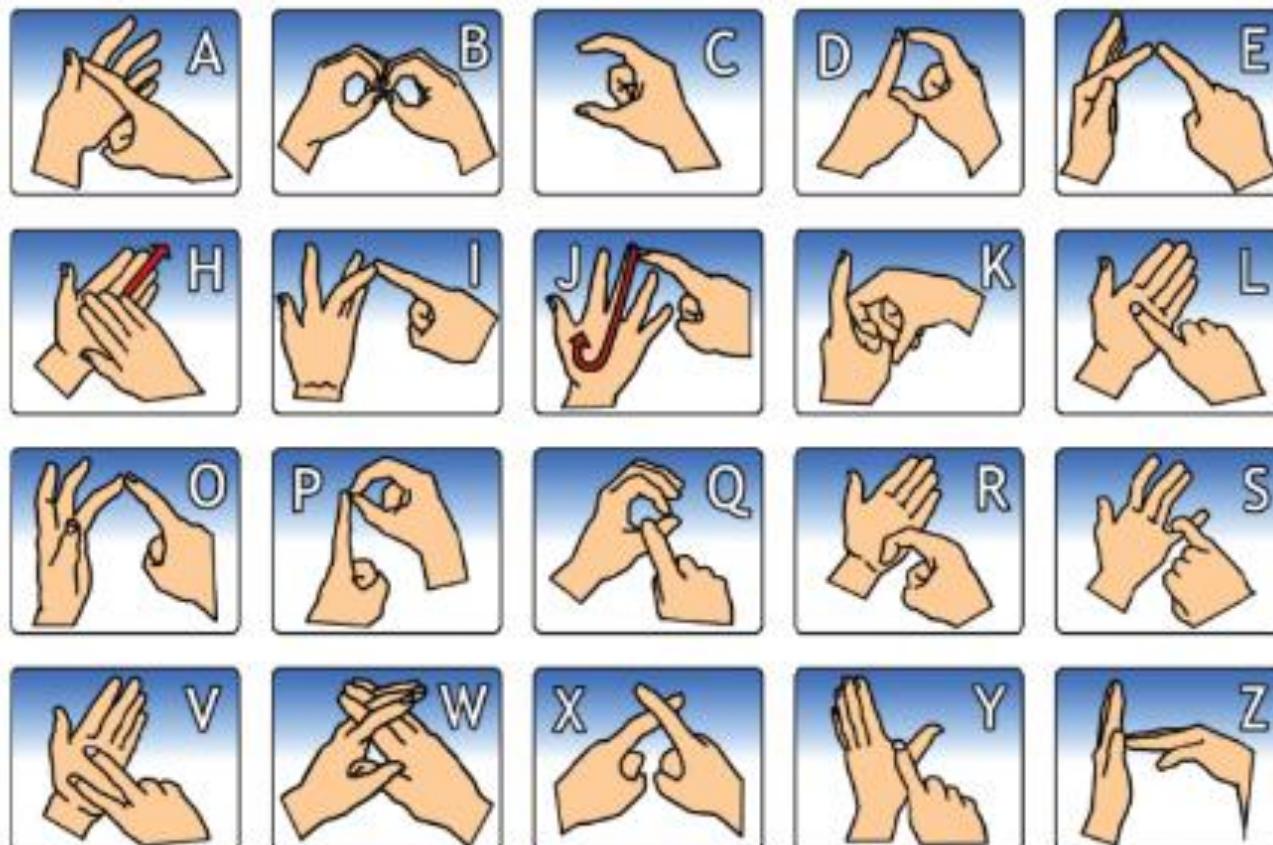
***Sign Languages are not universal.***

## American Sign Language (ASL)

1. Single handed Sign Language
2. Uses Object Subject Verb sentence structure
3. Contains complete vocabulary and grammar but expressed in physical movements of hands and arms.
4. Even within ASL, there are both “slang” movements and slight regional distinctions in how letters and words are formed.
5. 9 out of 10 children born deaf are born to parents who can hear.  
Still, even if parents are not initially fluent in sign language, their children often learn ASL fluently due to the incredible ability of young children to learn languages.
6. Behind English, Spanish and Chinese, ASL is the fourth most widely used language in the US with around 2 million users.



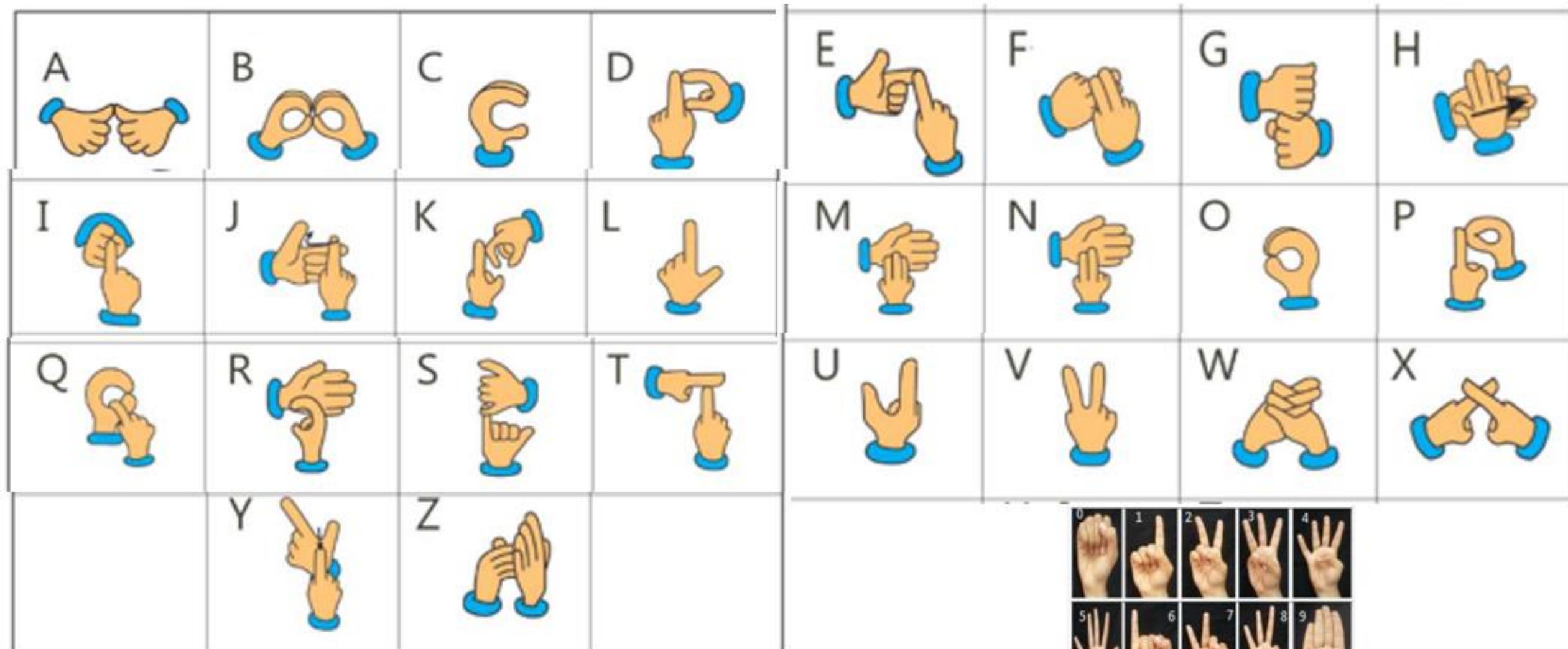
# British Sign Language (BSL)



**RIGHT  
HANDED**



# Indian Sign Language (ISL)





## Indian Sign Language Research and Training Centre

Department of Empowerment of Persons with Disabilities (Divyangjan)

Ministry of Social Justice & Empowerment, Government of India



<https://islrtc.nic.in/> website has ISL Dictionary which has 10,000 words that are represented in Indian Sign Language.

These 10,000 words are present in <https://disabilityaffairs.gov.in/islrtc/#> and divided into different sub-groups:

1. Academic Terms
2. Agriculture words
3. Civil Engineering Words
4. Computer Word List
5. Electrical and Electronics Words
6. Everyday Terms
7. Fashion Designing Words
8. Legal Terms
9. Medical Word List
10. Technical Terms
11. Tools and Equipment Word List

Every word of ISL dictionary is accompanied with Sign Language video which is too lengthy and difficult to remember the action sequence. Complex words like “wheeled robot” is also demonstrated in Sign Language.

Online Basic ISL course in Self Learning mode is there at

<https://www.youtube.com/playlist?list=PLFjydPMg4DapfRTBMokI09Ht-fhMOAYf6>

A thumbnail image from a YouTube playlist for "ISLRTC NEW DELHI". It shows a man in a dark shirt signing. Below the video, the title reads "Online Basic Indian Sign Language Course in Self Learnin..." and the channel name is "ISLRTC NEW DELHI".

Online Basic Indian Sign Language Course in Self Learnin...

ISLRTC NEW DELHI

41 videos 597 views Updated 2 days ago

⋮

Play all Shuffle

ISLRTC has developed a Basic Communication Skills in Indian Sign Language course, designed to be used in online self-learning mode. This course was formally launched by the Hon'ble State Minister of Social Justice and Empowerment on 23rd September 2023 on the occasion of Sign Language Day. The primary aim of this course is to foster basic communication skills in Indian Sign Language.

- 1 Guideline  
ISLRTC NEW DELHI • 352 views • 6 days ago
- 2 Module 1.1 Manners and etiquettes  
ISLRTC NEW DELHI • 218 views • 6 days ago
- 3 Module 1.2 Greeting and salutations in Indian Sign Language  
ISLRTC NEW DELHI • 331 views • 6 days ago
- 4 Module 1.3 Some polite useful phrases.  
ISLRTC NEW DELHI • 178 views • 6 days ago
- 5 Module 2 .1 Understanding the Indian sign language manual alphabet.  
ISLRTC NEW DELHI • 66 views • 6 days ago
- 6 Module 2.2 Introducing oneself (name, work, place, etc.)  
ISLRTC NEW DELHI • 75 views • 6 days ago

ISL E-CONTENT  
OF NCERT TEXTBOOKS  
developed jointly by  
ISLRTC & NCERT  
for classes I – V.  
The content is at  
<https://www.diksha.gov.in/>

The website contains Audio  
and interactive lessons for  
NCERT textbooks along with  
Sign Language Videos

(NEW) Marigold  
CBSE • Hindi,English • Class 4

All Video

WAKE UP !

WAKE UP...!

Sign Language Videos

Wake Up

Wake Up

Credits and Licence information

A screenshot of the ISL E-Content platform. At the top, it shows a sign language video titled "Wake Up" featuring a man in a pink shirt. To the right, there is a text box with a poem:  
Wash and dress  
And come on out –  
Everyone is up and about.  
The cow, the horses,  
the ducks and the sheep,  
The tiniest chicken  
Cheep-cheep-cheep  
Wake up!  
– C. Fletcher  
A small illustration of a girl running with a dog follows the text. A QR code is located in the bottom right corner of the video area.

English Marigold 2  
CBSE • English • Class 2

in

out

Toondemy- Opposites (The wind and the sun)

A screenshot of a Toondemy Opposites game for Class 2. It features two wooden houses with cats inside. Below each house is a button labeled "in" and "out". A hand cursor is pointing at the "out" button under the second house. The background is a green outdoor scene with trees and a path.

English Marigold 2  
CBSE • English • Class 2

cold

hot

Toondemy- Opposites (The wind and the sun)

A screenshot of another Toondemy Opposites game for Class 2. It shows a campfire on the left and ice cubes on the right. Below each image is a button labeled "cold" and "hot". A hand cursor is pointing at the "hot" button under the campfire. The background is a green outdoor scene with trees and a path.

English Marigold 2  
CBSE • English • Class 2

Toondemy- Opposites (The wind and the sun)

A screenshot of a third Toondemy Opposites game for Class 2. It shows a Santa Claus figure sledding down a snowy hill. A hand cursor is pointing at the sled. The background is a snowy winter scene with trees and falling snow.





## Deaf Online Education E-Learning

@DeafOnlineEducationELearning 18.7K subscribers 263 videos

This channel is related to Indian Sign Language education for deaf student... >

[facebook.com/deafonlineeducation](https://facebook.com/deafonlineeducation) and 4 more links

 Subscribed

HOME

VIDEOS

SHORTS

PLAYLISTS

COMMUNITY

CHANNELS

ABOUT



Latest

Popular

Oldest



INTERNATIONAL WEEK OF THE DEAF 18th to 24th September 2023| Nav Vani School...



how to know reading English sentences?

7.1K views • 5 months ago



Keep on continuing DOE E-Learning!

1.1K views • 11 months ago



Sign Week of the days & Year of the months

4.3K views • 11 months ago

# Sign Learn

ISLRTC

4.5★  
112 reviews10K+  
Downloads3+  
Rated for 3+ ⓘ[Install](#)[Share](#)[Add to wishlist](#)

This app is available for all of your devices

The screenshots show the following sequence:

- Home screen: "Choose Your Language" with a dropdown set to "ENGLISH". A "Next >" button is at the bottom right.
- Category selection screen: "Choose a Category" with options: Academic, Agriculture, Countries, Idiom, Legal, Medical, Numbers, and Regional.
- Search results screen: Search for "India" showing results like "higher education commission of india hec", "indian council of agricultural research icar", etc.
- Video player screen: "village sign language" showing a video of a person signing. Buttons for "share" and "Favourites" are visible.
- Home screen: "Choose Your Language" with a dropdown set to "ENGLISH". A "Next >" button is at the bottom right.
- Category selection screen: "Choose a Category" with options: Academic, Agriculture, Countries, Idiom, Legal, Medical, Numbers, and Regional. A circular arrow icon is overlaid on the "Idiom" button.

# Real problem

**All the resources of Sign Language are using Sign Language ranging from alphabets to Text Books.  
Making people depend on Signs throughout their life and separating them from normal people**

- When normal people see any new word they refer dictionary and understand the meaning without sound
- After acquiring sufficient instructions and knowledge using sound, sound is no more necessary for learning
- Reading and Writing skills help to do text-chatting on devices by any person who can see can communicate
- With the advent of graphics and videos tagged with texts, learning visible-equivalents of sound-instructions of childhood can be possible

**How much knowledge is necessary to learn English language?**

- Plain English is a mode of writing or speaking the English language intended to be easy to understand regardless of one's familiarity with a given topic. Plain English removes ambiguity and so became official.
- Basic English is a controlled language based on standard English, but with a greatly simplified vocabulary and grammar. It was created by the linguist and philosopher Charles Kay Ogden as an international auxiliary language, and as an aid for teaching English as a second language. It contains only 850 words
- Oxford Advanced Learner's dictionary is composed of using 3500 keywords. Any meaning can be expressed using these keywords. The dictionary contains 1,50,000 thousand words explained by these keywords. Of 3500 keywords major words are nouns, verbs and adjectives. Structural words are only 55.
- By extracting basic essential words and minimum grammatical structures any one can communicate by reading and writing and learn further any domain, based on this basic knowledge.

| TITLE & JOURNAL   | PROPOSED TECHNIQUES AND ALGORITHMS   | DATASET   | PREPROCESSING FACTORS   | PERFORMANCE FACTORS AND MODELS COMPARED | RESULTS ACHIEVED  | POSITIVES  | NEGATIVES   |
|---|--|---|---|---|---|--|---|
| <b>A Modified-LSTM Model for Continuous Sign Language Recognition using Leap motion, IEEE Xplore</b>  | Continuous SLR using Leap motion sensor. Splitting of continuous signs into sub-units and modeling them with neural networks, Modifield-LSTM | A dataset of 35 isolated sign words has been used while training the model. evaluation done on 942 signed sentences produced by six signers   | Redundant information such as, joints of fingers and wrist are removed. Only 12 features of hands are taken and size normalization is done between [-1,1] | Traditional LSTM                        | The average accuracy of 72.3% and 89.5% have been recorded on signed sentences and isolated sign words, respectively. | Indian Sign Language. Dataset size is lesser than traditional LSTM model | Applied on 35 symbols only, less symbols. Used motion |
| <b>Real-Time AmericanSign Language Recognition Using Desk and Wearable Computer Based Video, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE</b> | wearable computer system for sentence-level continuous ASL, Hidden Markov models   | Selected pronouns, verbs, nouns and adjectives NTSC composite video is captured and analyzed at 320 ¥ 243 pixel resolution. A Silicon Graphics 200MHz R4400 Indy workstation maintains hand tracking at 10 frames per second, | Outliers are removed  |   | Desk based system with 92 percent word accuracy and cap worn system with 98 percent accuracy                          | Higher accuracy  | Carrying devices                                      |

| TITLE & JOURNAL  | PROPOSED TECHNIQUES AND ALGORITHMS   | DATASET   | PREPROCESSING FACTORS  | PERFORMANCE FACTORS AND MODELS COMPARED  | RESULTS ACHIEVED  | POSITIVES   | NEGATIVES        |
|--|--|---|--|--|---|---|------------------|
| <b>Development of an End-to-End Deep Learning Framework for Sign Language Recognition, Translation, and Video Generation , IEEE ACCESS</b> | Mediapipe was used to capture pose sequences.<br>VGG-19 model to classify<br>Bi-LSTM network for text generation | The model is trained using a dataset of 40,000 videos for 320 classes<br><br>Each class is recorded with 50 repetitions<br><br>RWTH-PHOENIX-Weather 2014T dataset: deals with the SLRT research for German sign language.<br>ISL-CSLTR dataset: SLRT research in Indian sign language<br><br>The How2Sign dataset SLRT research for American Sign Language. | The CNN network processes images of size $254 \times 254$ . It uses $3 \times 3$ filters with stride level 1. The max pooling operation is performed using stride level 2 and a window size of $2 \times 2$ . After this process, the dimensions of pixels are reduced to $112 \times 112 \times 64$ . Further, the convolution layer of varying filter size 128, 56, 28 is applied and reduces the size of the image as well as focuses the important features<br><br>After completion of preprocessing steps, the videos of high resolution to be $1920 \times 1080$ and | CNN+SVM<br><br>DeepLabv3+Bi-LSTM<br><br>LSTM+KNN<br><br>CNN+Bi-LSTM with attention<br><br>GAN+LSTM+3DCNN | The proposed model has achieved the above 95% classification accuracy towards SL recognition, 38.56 average BLEU score, remarkable human evaluation scores, 3.46 average FID2vid score, 0.921 average SSIM values, 8.4 average Inception Score, 29.73 average PSNR score, 14.06 average FID score, and an average 0.715 TCM Score | Achieved higher accuracy using hybrid and bi-LSTM<br><br>Achieved text to video animations faster with lesser dataset<br><br>Translation of multilingual - signlanguage | It is very slow. |

| TITLE & JOURNAL  | PROPOSED TECHNIQUES AND ALGORITHMS   | DATASET   | PREPROCESSING FACTORS   | PERFORMANCE FACTORS AND MODELS COMPARED   | RESULTS ACHIEVED   | POSITIVES   | NEGATIVES  |
|--|--|---|---|---|--|---|--|
| <b>A Prototype for Mexican Sign Language Recognition and Synthesis in Support of a Primary Care Physician, IEEE ACCESS</b> | <p>Four sets of data were obtained to determine which sign was made: sign trajectory, image contour, body area where the sign was made, and the number of hands involved in the sign performance</p> <p>Use a hidden Markov model (HMM) to classify the hand shape involved in the sign based on the X and Y coordinates of the contour points.</p> <p>Use another HMM to classify the sign based on the X, Y, and Z coordinates of the centroid and the left and right hands describing the signal movement</p> <p>For text to sign --- make intermediary</p> | 22 volunteers practiced 82 signs, five times each indoors. The Kinect sensor was located at a height of 1.3 meters and at 1.5 meters from the person. | <ul style="list-style-type: none"><li>Capture the images in RGB color.</li><li>· Binarize the images to detect the blobs.</li><li>· Represent the hand posture contour as a set of X and Y coordinates.</li><li>· Normalize the coordinates in a range between 0 and 1.</li><li>· Smooth the coordinates corresponding to the signal to reduce noise.</li></ul> | <p>Other Kinect models but with :</p> <ul style="list-style-type: none"><li>1.limited signers</li><li>2.controlled background</li><li>3.dataset without any particular semantic order</li></ul> | <p>accuracy and F1 scores average rates of 99% and 88%, a mean specificity rate of 99.80% was obtained for the general medicine consultation service words</p> | <ul style="list-style-type: none"><li>1.Animated videos for text</li><li>2.Very high accuracy</li><li>3.Two way communication</li></ul> | <ul style="list-style-type: none"><li>1.ONLY Indoors</li><li>2.Continuous use is not recommended as device heats and cannot give accurate results anymore.</li></ul> |

| TITLE & JOURNAL   | PROPOSED TECHNIQUES AND ALGORITHMS  | DATASET     | PREPROCESSING FACTORS   | PERFORMANCE FACTORS AND MODELS COMPARED  | RESULTS ACHIEVED   | POSITIVES   | NEGATIVES                            |
|---|---|-------------|---|--|--|---|--------------------------------------|
| <b>Indian Sign Language Recognition through Hybrid ConvNet-LSTM Networks,</b><br><b>EMITTER International Journal of Engineering Technology</b> | Spatial Video Sequences are handled by CNN and temporal Video Sequences are handled by RNN, Convolutional Neural Network and Recurrent Neural Network | CasTalk-ISL | Redundant information such as, joints of fingers and wrist are removed. Only 12 features of hands are taken and size normalization is done between [-1,1] | 1. video sequence with BLSTM-3D residual networks<br>2. Gesture Recognition Based on Kinect v2 and Leap Motion Data Fusion<br>3. 3D Convolutional Neural Networks for dynamic sign language recognition<br>4. Multi-modality-based Arabic Sign Language recognition,<br>5. Real-time sign language gesture (word) recognition from video sequences using CNN and RNN<br>6. Sign language recognition using Convolutional Neural Networks | 95.99% top-1 accuracy and 99.46% top-3 accuracy on the test dataset. | Word and Gestures are in sync with ISL research and Training center | Preprocessing stage is not discussed |

| TITLE & JOURNAL  | PROPOSED TECHNIQUES AND ALGORITHMS                                     | DATASET  | PREPROCESSING FACTORS   | PERFORMANCE FACTORS AND MODELS COMPARED                                    | RESULTS ACHIEVED  | POSITIVES  | NEGATIVES  |
|--|--|--|---|--|---|--|--|
| <b>Helping Hearing-Impaired in Emergency Situations: A Deep Learning-Based Approach, IEEE ACCESS</b> | Object detection using You Only Look Once (YOLO, version 8) algorithm. | ISL Signs for accident, call, doctor, help, hot, police, pain and thief. The dataset has 412 videos. The videos shot on 26 adult individuals 12 male and 14 female who are in the age between 22 to 26 | The videos are stored as image sequence and each image is resized from 500 X 600 to 150 X150 pixels. Data normalization is also done. | 3D CNN and a combination of pre-trained CNN and LSTM.<br>A multi-class SVM | YOLO successfully detected the Signs with good precision and overall recall score 99.5% | YOLO uses a single CNN for both classification and localization.<br>YOLO is fast with processing speed of 40 to 50 images per second | YOLO is best used for static Signs but not for Action Signs. |

| TITLE & JOURNAL  | PROPOSED TECHNIQUES AND ALGORITHMS  | DATASET   | PREPROCESSING FACTORS  | PERFORMANCE FACTORS AND MODELS COMPARED | RESULTS ACHIEVED  | POSITIVES  | NEGATIVES   |
|--|---|---|--|---|---|--|---|
| <p><b>Object Classification Using Sequences of Zernike Moments, HAL Open Science, 16th IFIP International Conference on Computer Information Systems and Industrial Management (CISIM), Jun 2017, Bialystok, Poland.</b></p> | <p>1. Calculation of Zernike moment sequence for every image in the data set.<br/>2. Construction of the class prototype using Zernike moment sequence.<br/>3. Testing of image classification on the data set using the reference sequence.<br/>Zernike moments.<br/>MSE criterion for tolerance</p> | <p>1. MPEG-7 Core Experiment CE-Shape-1 Test Set from <a href="http://www.dabi.temple.edu/">http://www.dabi.temple.edu/</a></p> | <p>Images are converted to binary before applying Zernike Moments.</p> |   | <p>1. The animal shape image classification is the most mixed group as some of the results fall within the 0.85 – 1.00 accuracy range and some into the 0.5 – 0.6 accuracy range<br/>2. the 3D object image classification is in the 0.7 – 1.0 range<br/>3. The results for the macroscopic image classification falls into the range of 0.8 – 1.0<br/>4. The accuracy of classifying digits is less than 0.5</p> | <p>Images are classified irrespective of size, translation and rotation because of this Pre-processing steps are minimized</p> | <p>The accuracy of identifying digits is poor because of similarity in the image features extracted by Zernike moments.</p> |

# Problem Statement

## Sentence Identification -

1. Deep learning is used to train the model to display the sentence of the gesture done in ISL. If the accuracy value is lesser it displays repeat the gesture and if the accuracy is too low it displays can't understand.
2. It displays one of the trained sentences and asks us to perform the gesture. If it is correct it displays correct otherwise wrong.

## Sentence formation -

### A) Pre-defined sentence structure

1. Signs are performed in the sequence
2. Combining all the words to display the message

### B) Type and word pair

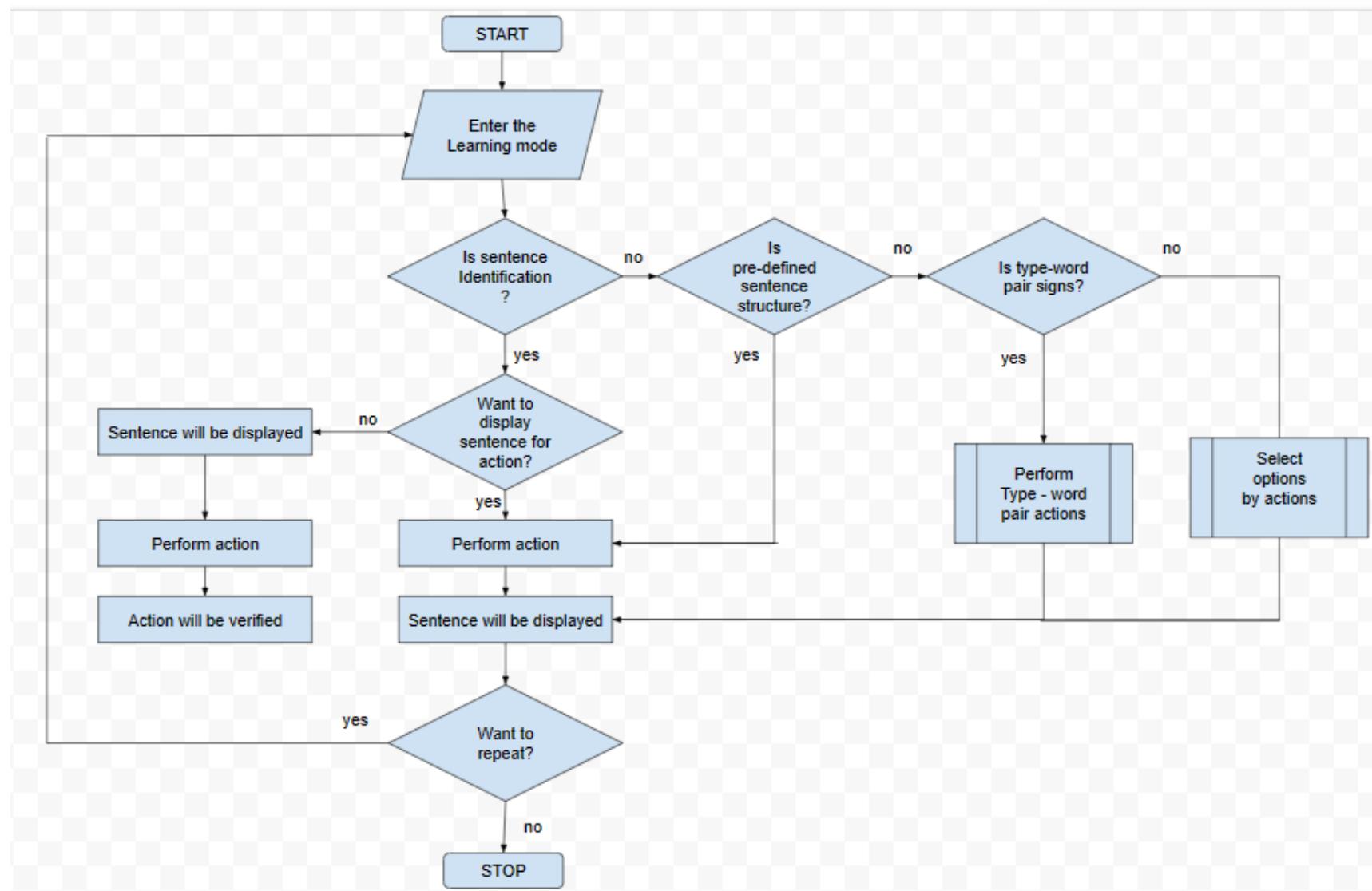
1. First Sign for selecting parts of speech
2. Second Sign for actual word
3. Repeating above for all the words
4. Combining the all words to display the message

### C) Selecting Options

1. Options with words are displayed
2. Sign is performed to select the word
3. Repeating above for all the words
4. Combining the all words to display the message

# Architecture

- Instead of training all signs in one model, division of words based on category is done
- Models will be trained for each sub-division
- Object detection models or Action detection models are used based on category
- The recognition-results of the sub-divisions will be combined and the resulting sentence will be displayed
- In second mode sentence is displayed. If we perform the corresponding action it will display whether it is correct or wrong.



# Technologies used



- Machine Learning
- Deep Learning
- Open Computer Vision (OpenCV)
- MediaPipe for innovative machine learning features

# Models done

## Deep Learning Models

- Convolution Neural Network (CNN)
- You Only Look Once(Yolov8)
- Recurrent Neural Network(RNN)
- Long Short Term Memory – Recurrent Neural Network (LSTM-RNN Hybrid model)
- Modified Gated Recurrent Unit(MOPGRU)

# References

22. Finger Counter using Hand Tracking | Computer Vision | OpenCV Python 2021  
<https://www.computervision.zone/courses/finger-counter/>

# Experimental Results

B A B A

A

B

C

B

A

B

C

B

A

C

how\_are\_you good\_morning hello good\_

hello

good\_morning

how\_are\_you

hello good\_morning how\_are\_you good\_

hello

good\_morning

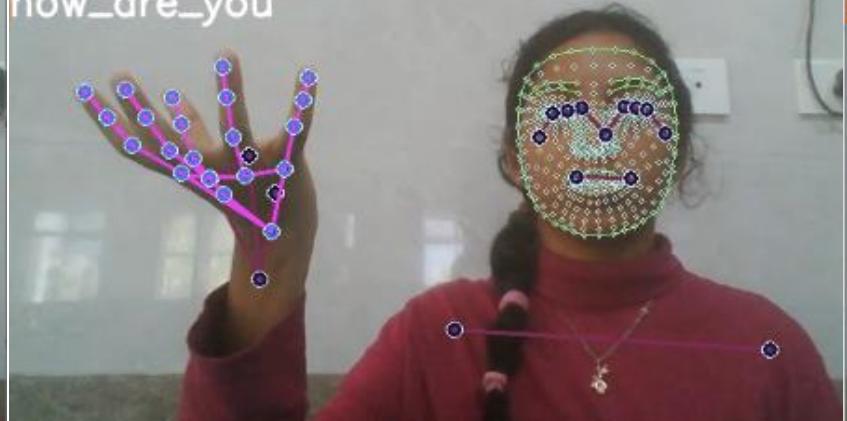
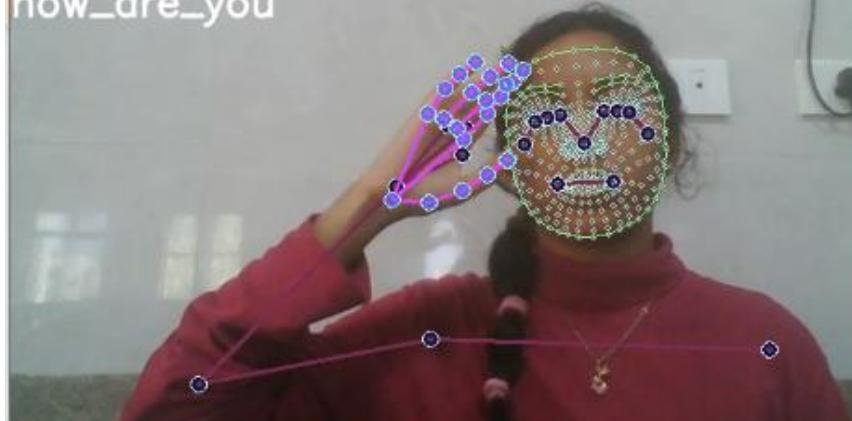
how\_are\_you

good\_morning hello good\_morning hello

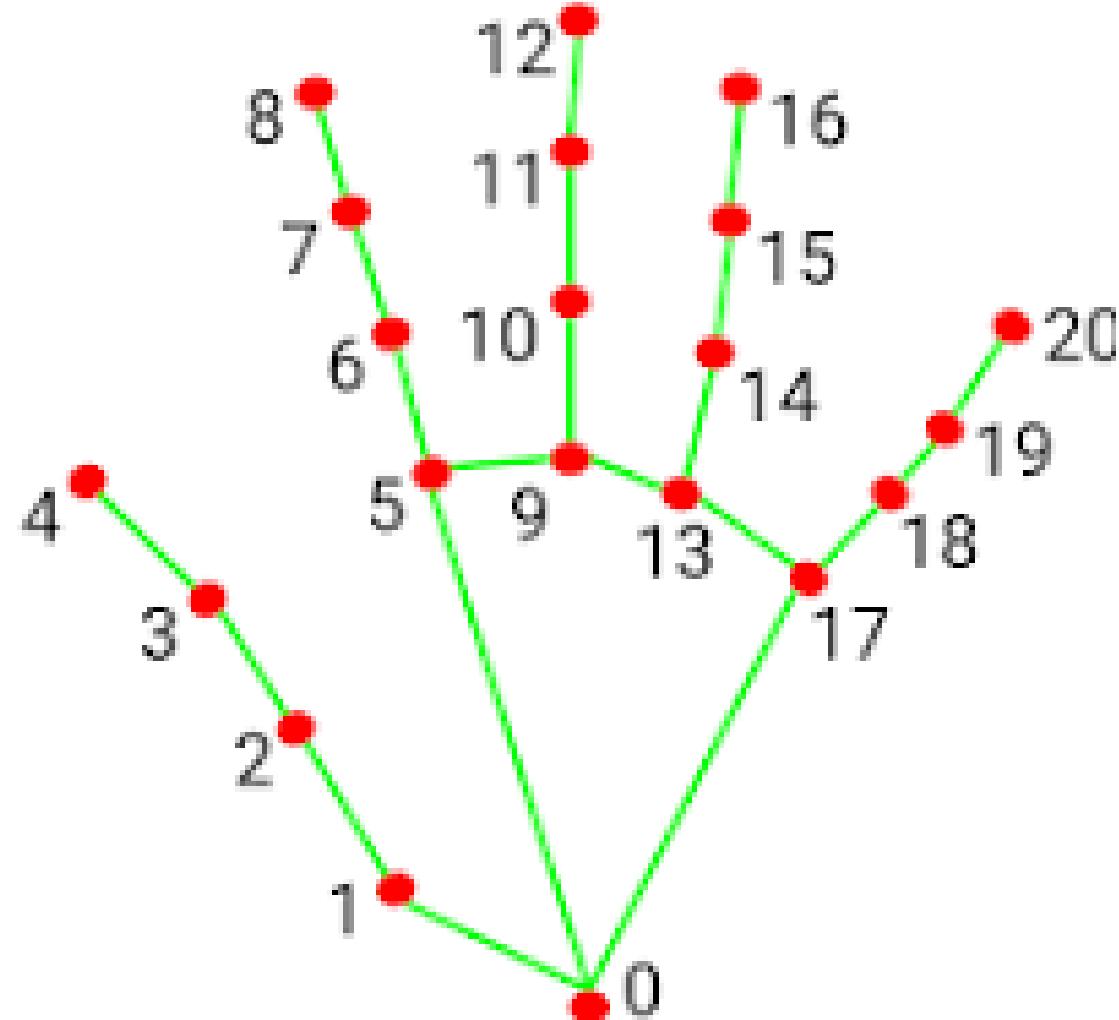
hello

good\_morning

how\_are\_you



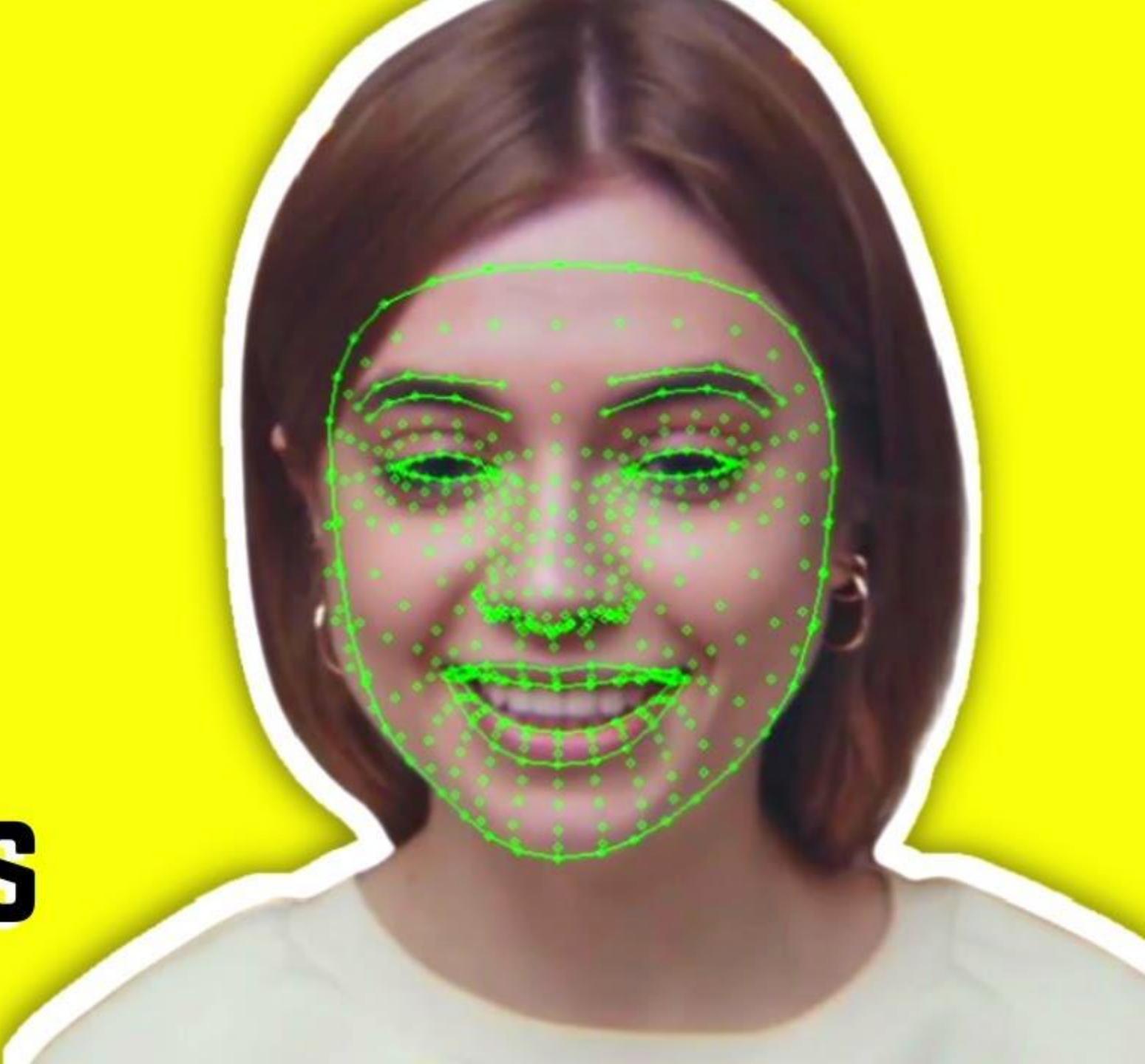
# Experimental Results



## Hand Landmarks

LATEST

**468**  
**FACE**  
**LANDMARKS**



# **WELCOME**

**1.DIGITS**

**2.LETTERS**

**9.EXIT**







C



I



L



O



U



V



1



DIGITS

2



LETTERS

3



NUMBERS

RELIANCE ENTERTAINMENT & PHANTOM FILMS PRESENT  
A KABIR KHAN FILMS PRODUCTION



4



WORDS

5



SYMBOLS

# INDIAN SIGN LANGUAGE

BY ISH NEWS

+

# 83

ISH  
NEWS

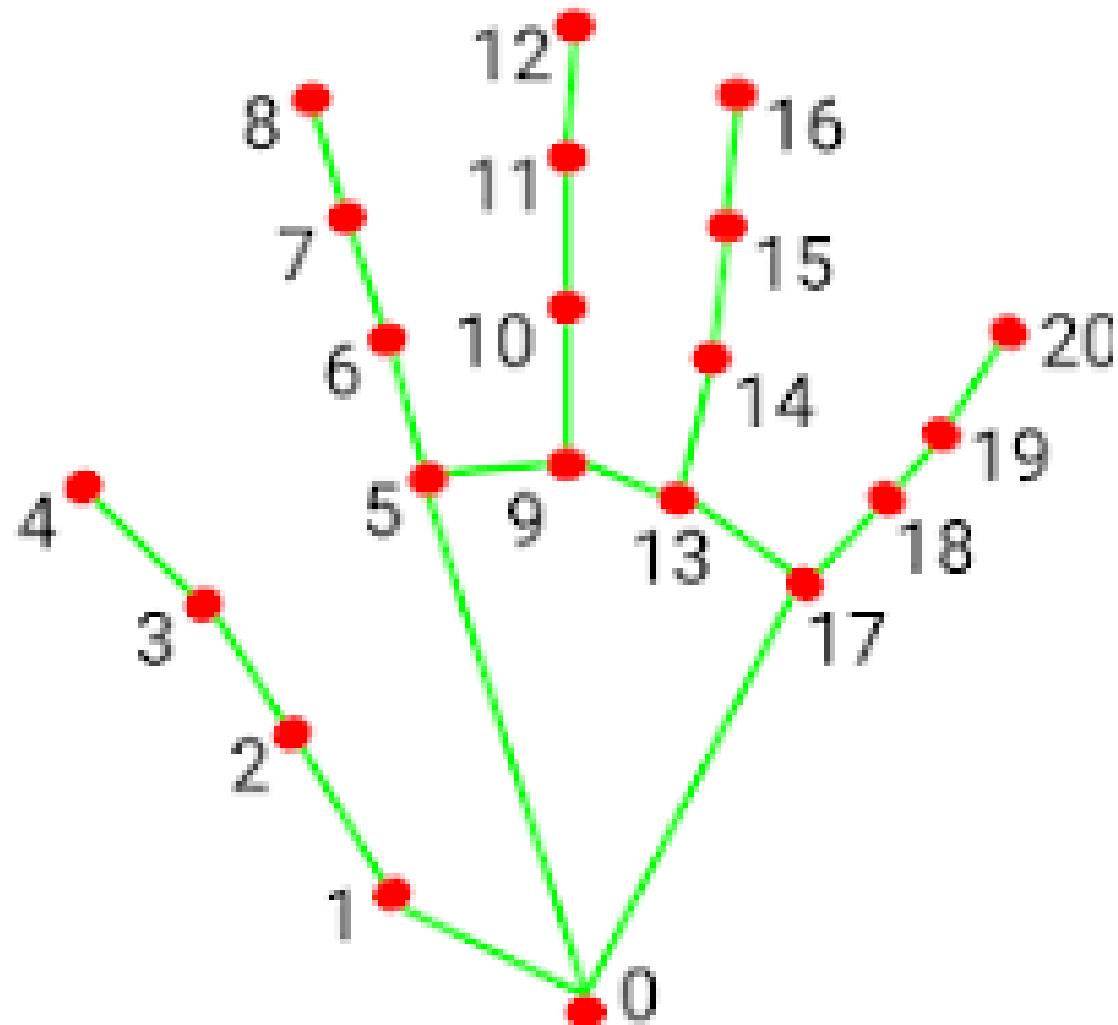


PRODUCED BY DEEPIKA PADUKONE, KABIR KHAN, VISHNU VARDHAN INDURI, SAJID NADIADWALA,  
SHEETAL VINOD TALWAR, RELIANCE ENTERTAINMENT & 83 FILM LTD



EXIT

# Left hand vs Right hand



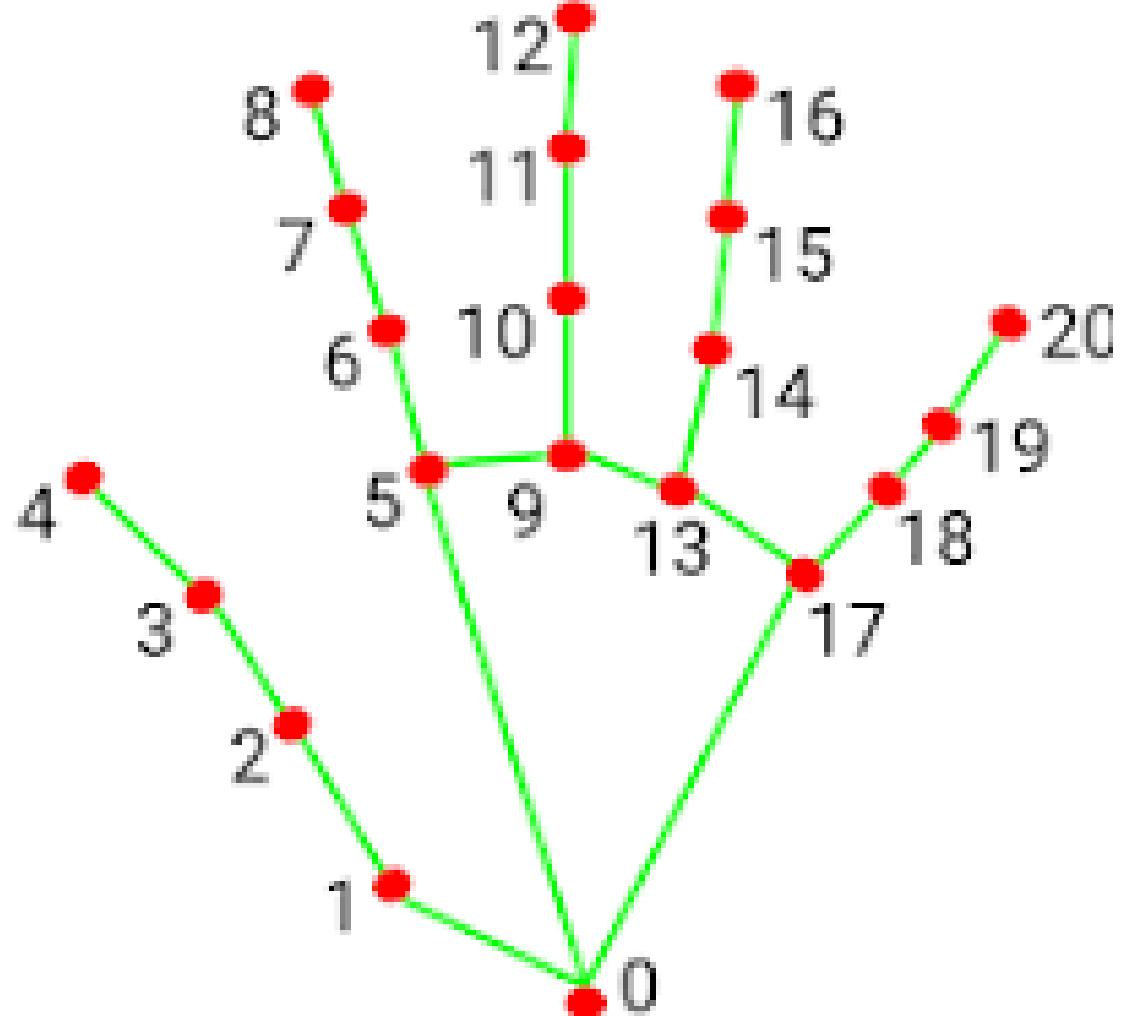
If X co-ordinate of 5<sup>th</sup> mark is lesser than X co-ordinate of 17<sup>th</sup> mark than it is left hand otherwise right hand.

Positive X direction is from left to right

Positive Y direction is from top to bottom

## Hand Landmarks

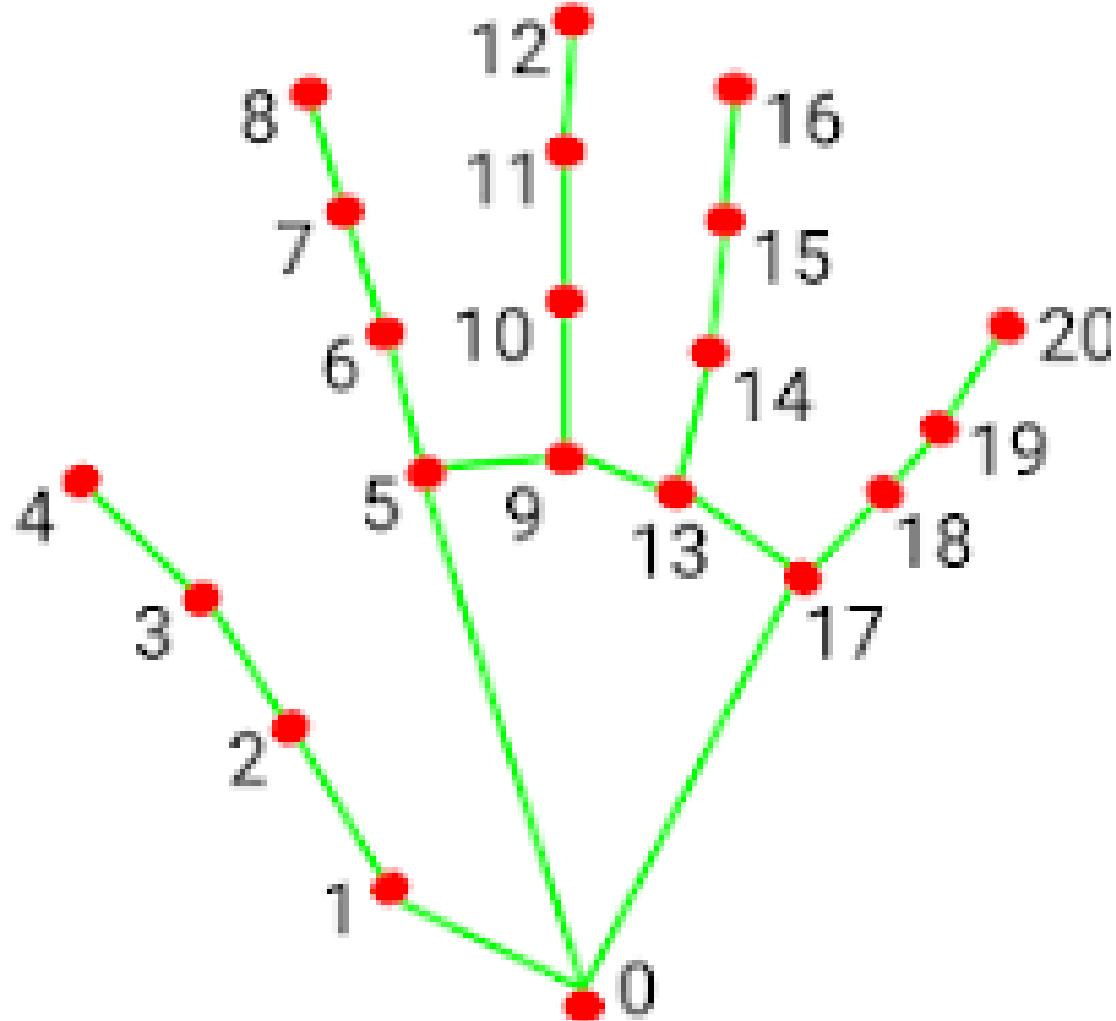
# Possibility of two Left hands or two Right hands



**As Hand Landmarks are not considering palmar side or dorsal side of the hand we can pose both hands as shown in the figure with left hand at palmar side and right hand with dorsal side and both of them are considered left hands**

## Hand Landmarks

# Z-axis co-ordinates



**MediaPipe gives us Hand Land Marks in all three directions. The positive direction of Z-axis is into the screen.**

## Hand Landmarks

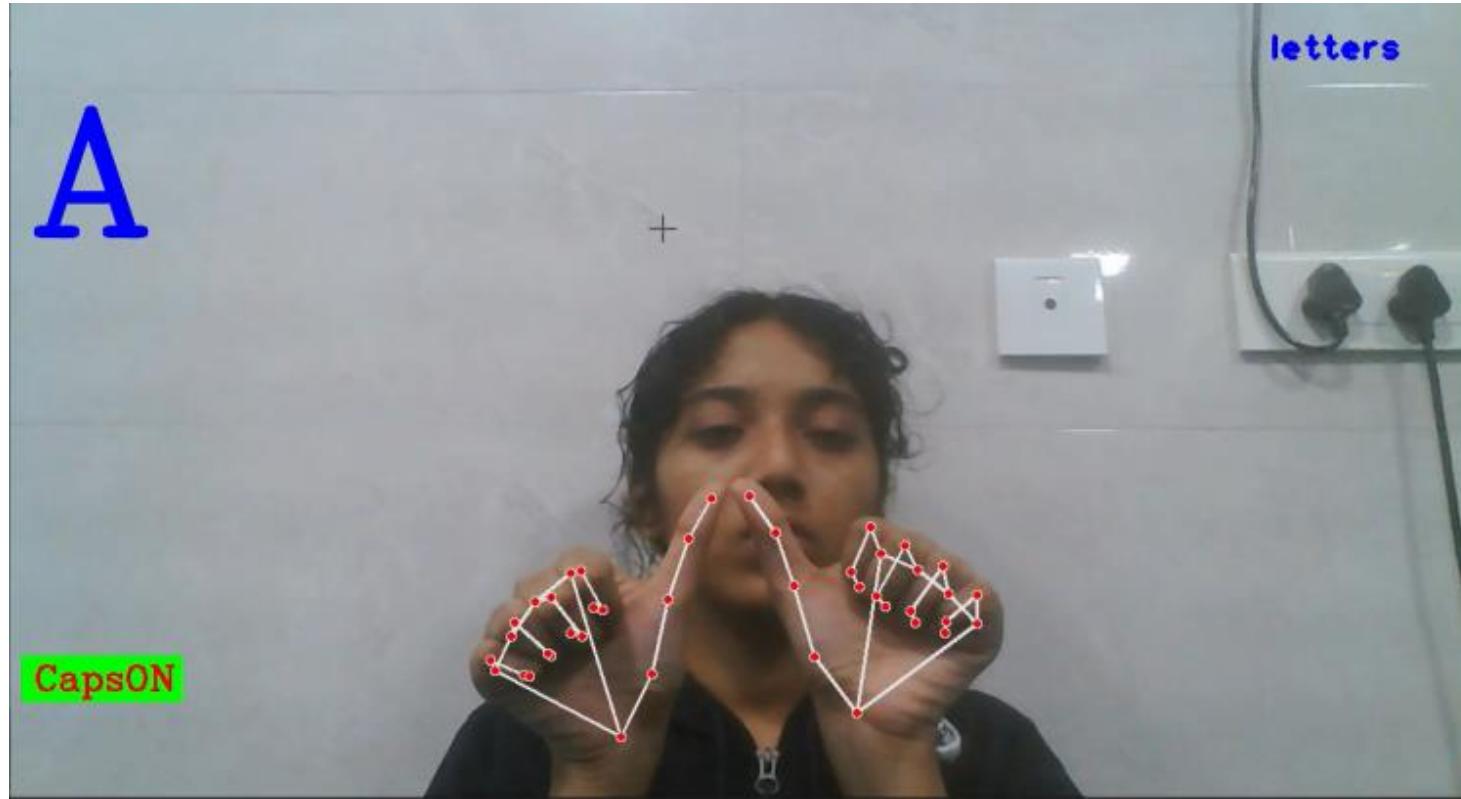
# Open finger vs Closed finger



## Example: Index finger

If Y co-ordinate of 8<sup>th</sup> mark is lesser than Y co-ordinate of 6<sup>th</sup> mark than the index finger is opened otherwise index finger is closed.

# Articulation of Letter 'A'



1. Index, middle, ring and litter fingers are closed for both left and right hands.
2. Thumbs are opened for both left and right hands
3. X and Y co-ordinates of 4<sup>th</sup> mark for both hands are very nearly close.

# Orientation of hands



1. Determination of opening and closing of fingers depends on Hand orientation.
2. Zero, 90, 180 and 270 degrees orientations are included in the project
3. In the above articulation for 'H' symbol is with left hand as left hand in 180 degree orientation and right hand as left hand with 90 degree orientation and all fingers are opened

# First hand - Second hand



**When we try to articulate the letter 'H' even if we show the left hand first it is not guaranteed that MediaPipe will give us HandLandmarks of left hand first, it may give the HandLandmarks of right hand first. So, both cases are considered in the project.**

A



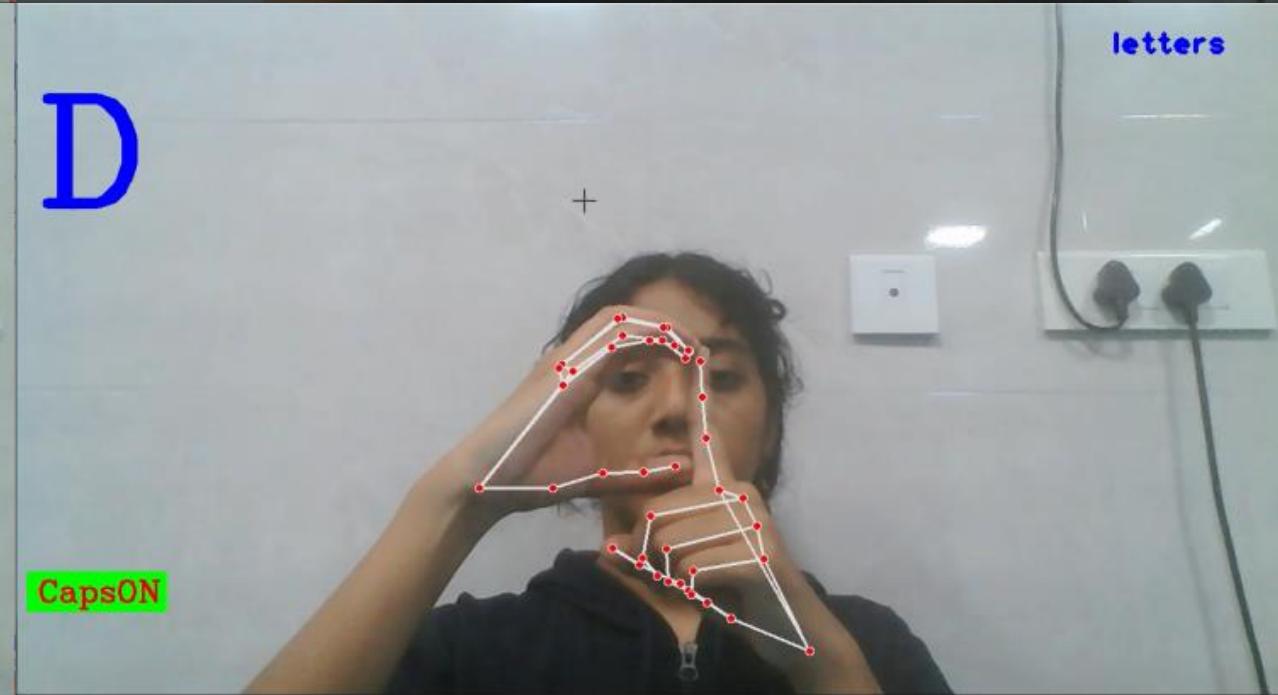
B



C



D



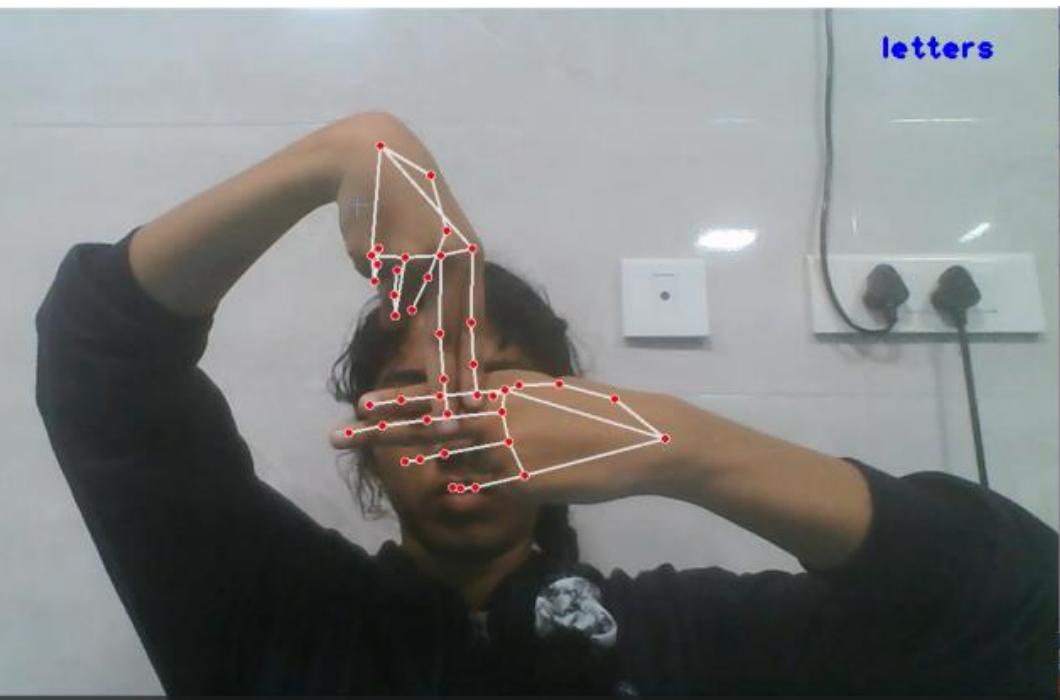
E

CapsON



F

CapsON



G

CapsON



H

CapsON



I

CapsON

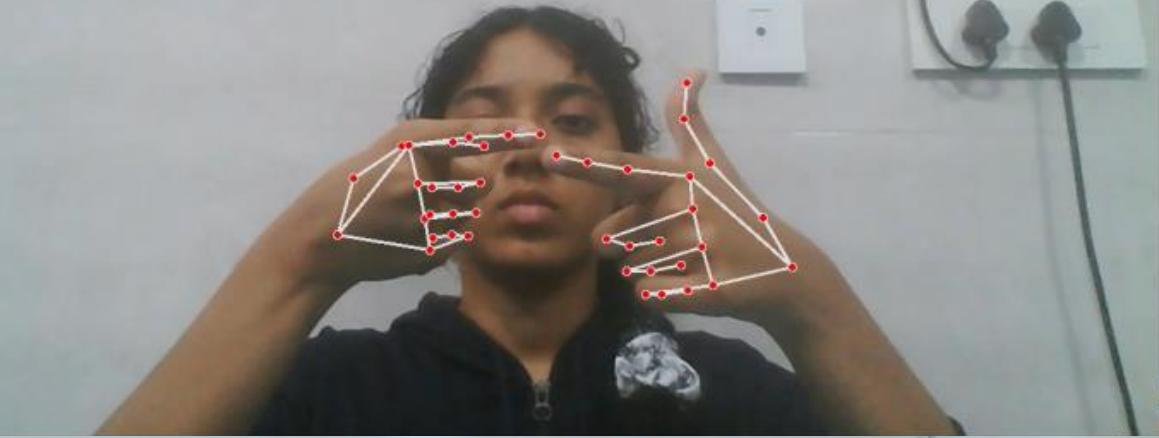


letters

j

+

letters



k

letters

l

+

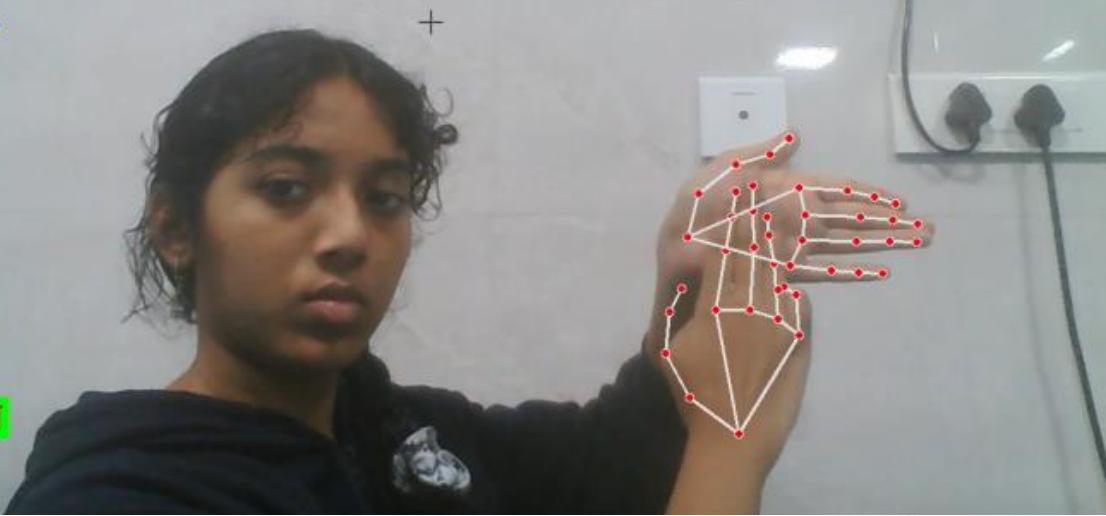
letters



# M

CapsON

letters



# N

CapsON

letters



# O

CapsON

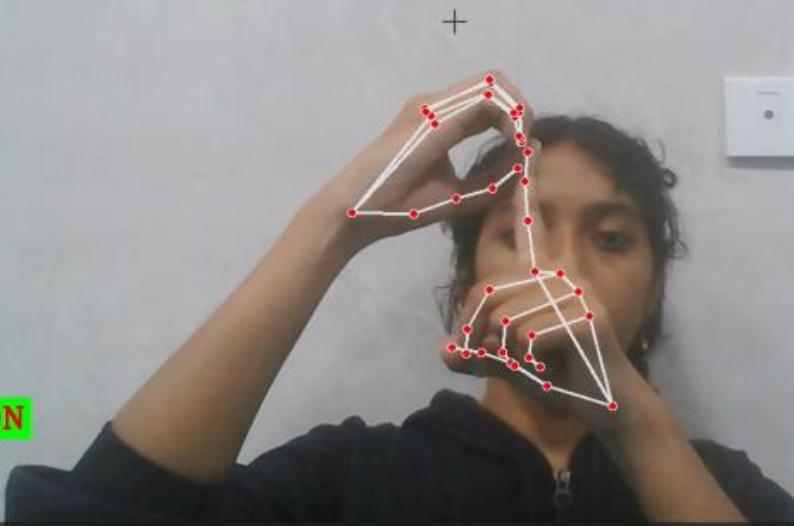
letters



# P

CapsON

letters



**Q**

CapsON

**S**

CapsON

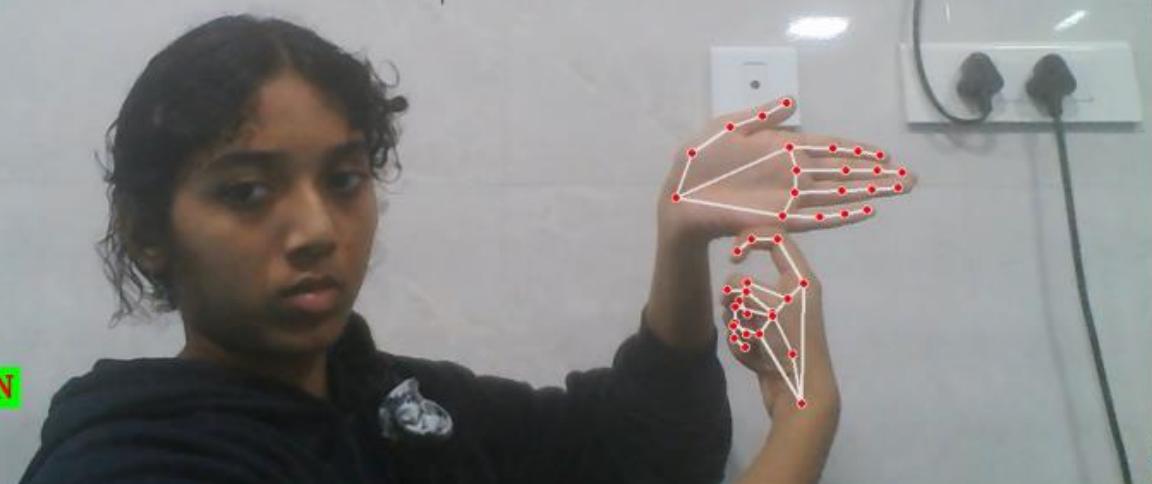
letters



**R**

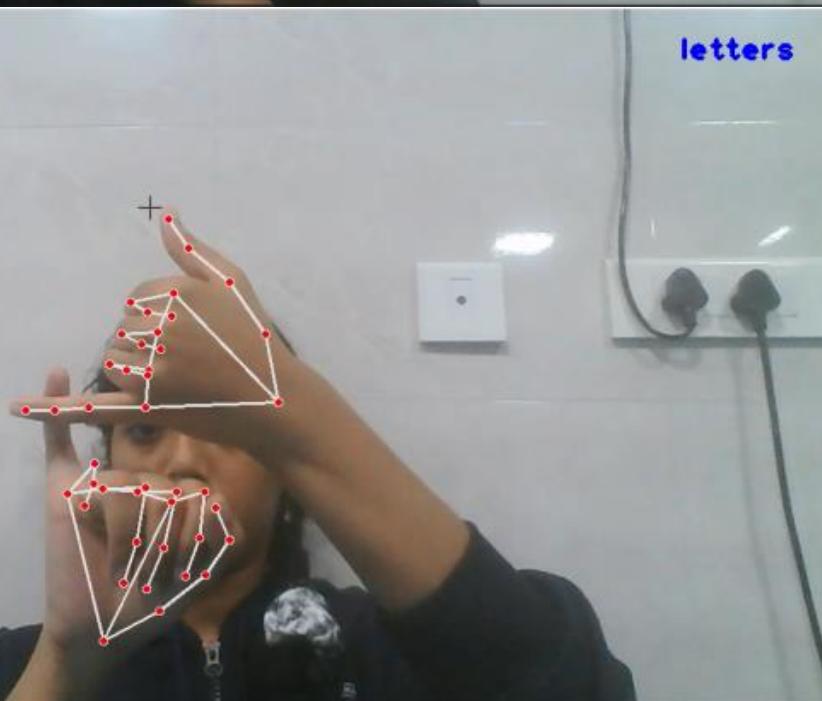
CapsON

letters



**T**

CapsON



# U



# V



# W



# X



# Y

CapsON

Choose a symbol

ABCDEFIGHJK

letters

+



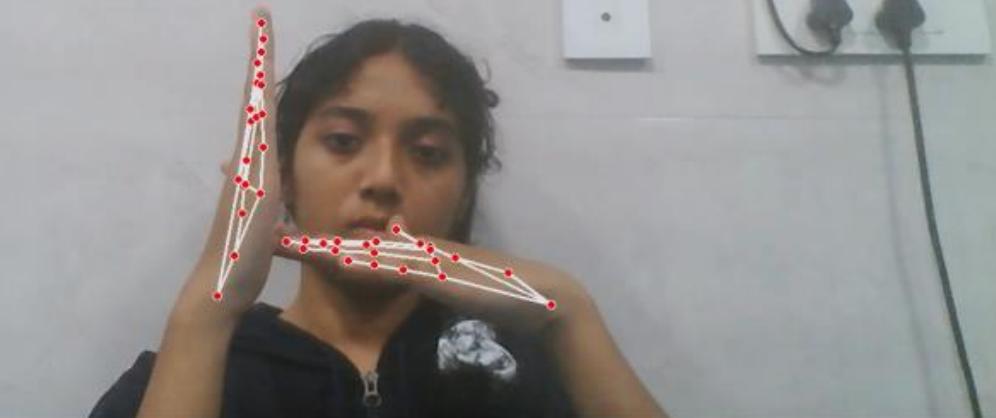
# Z

CapsON

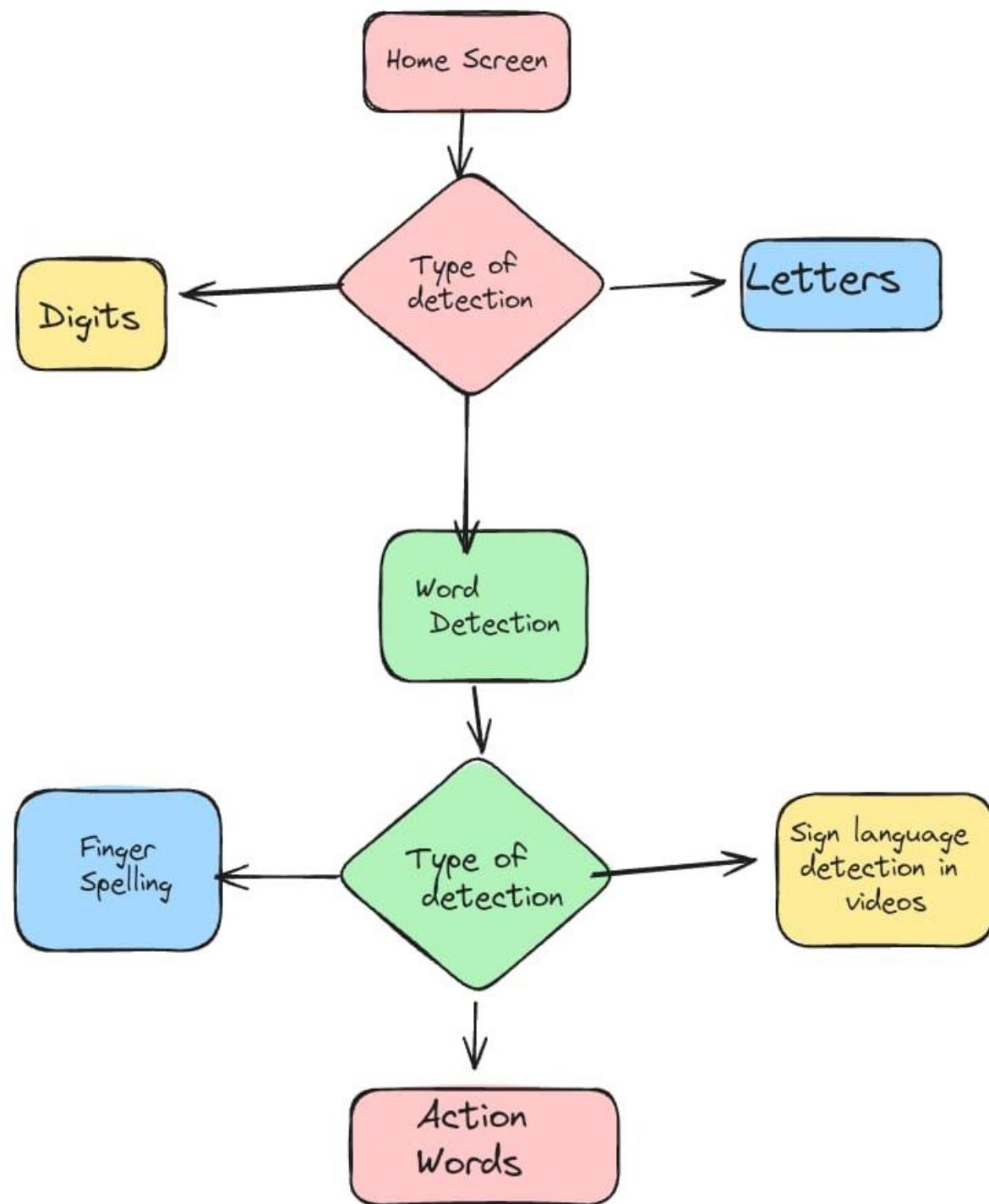
words

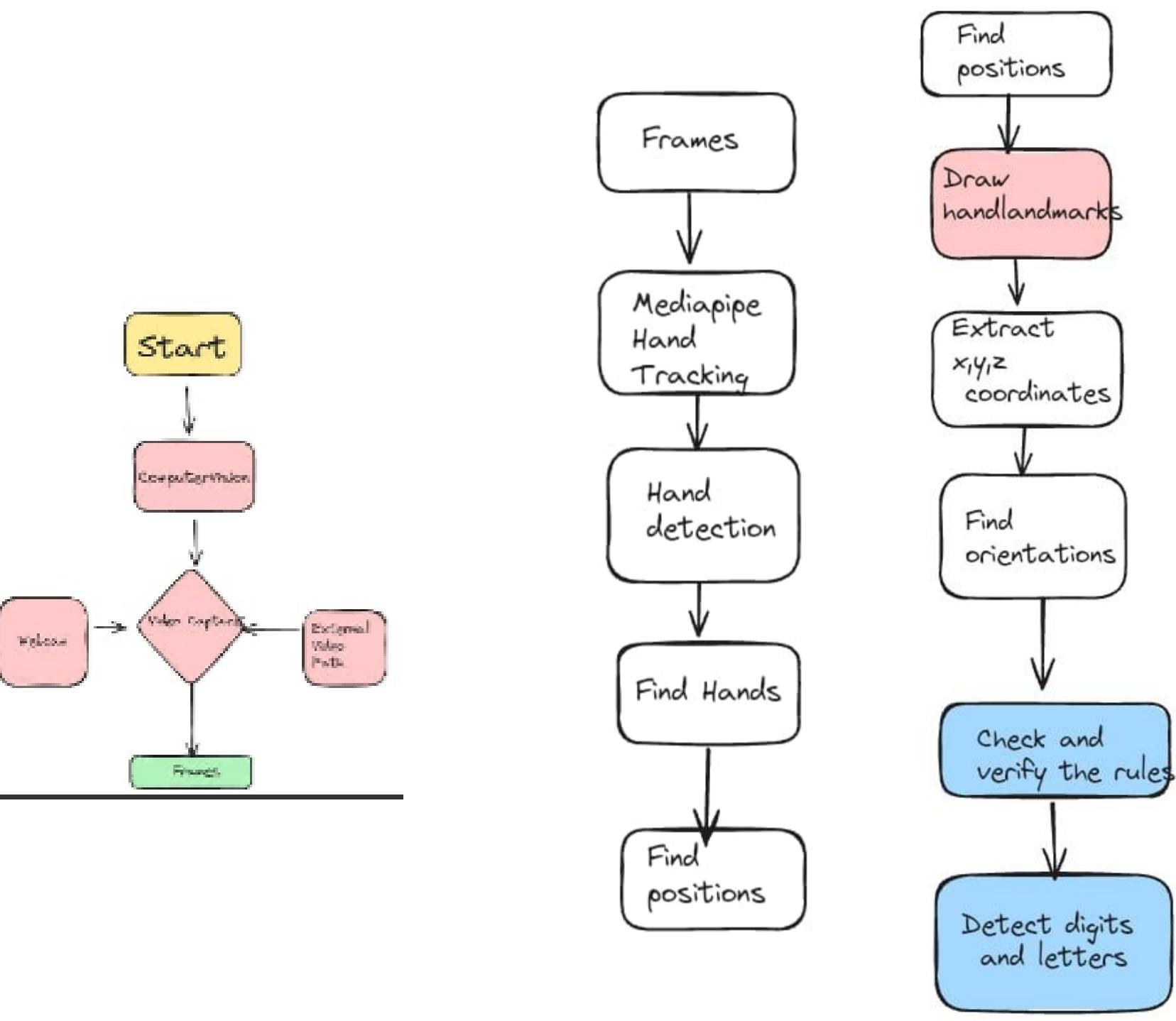
letters

+



# PHASE - 2





# Finding and determining the movements

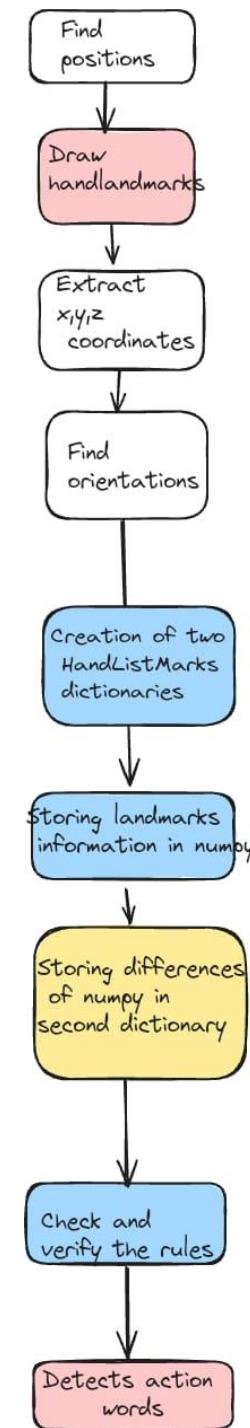
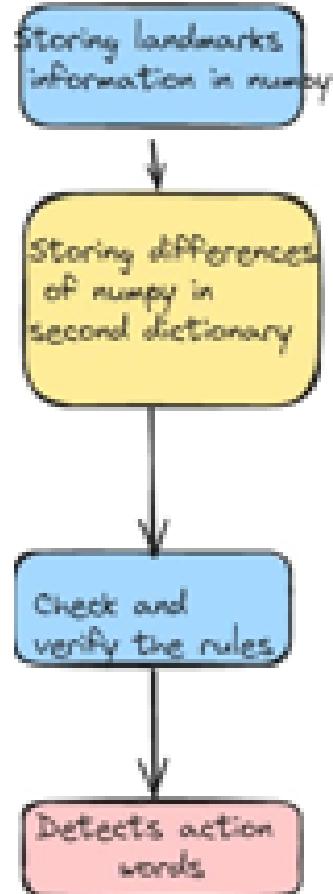
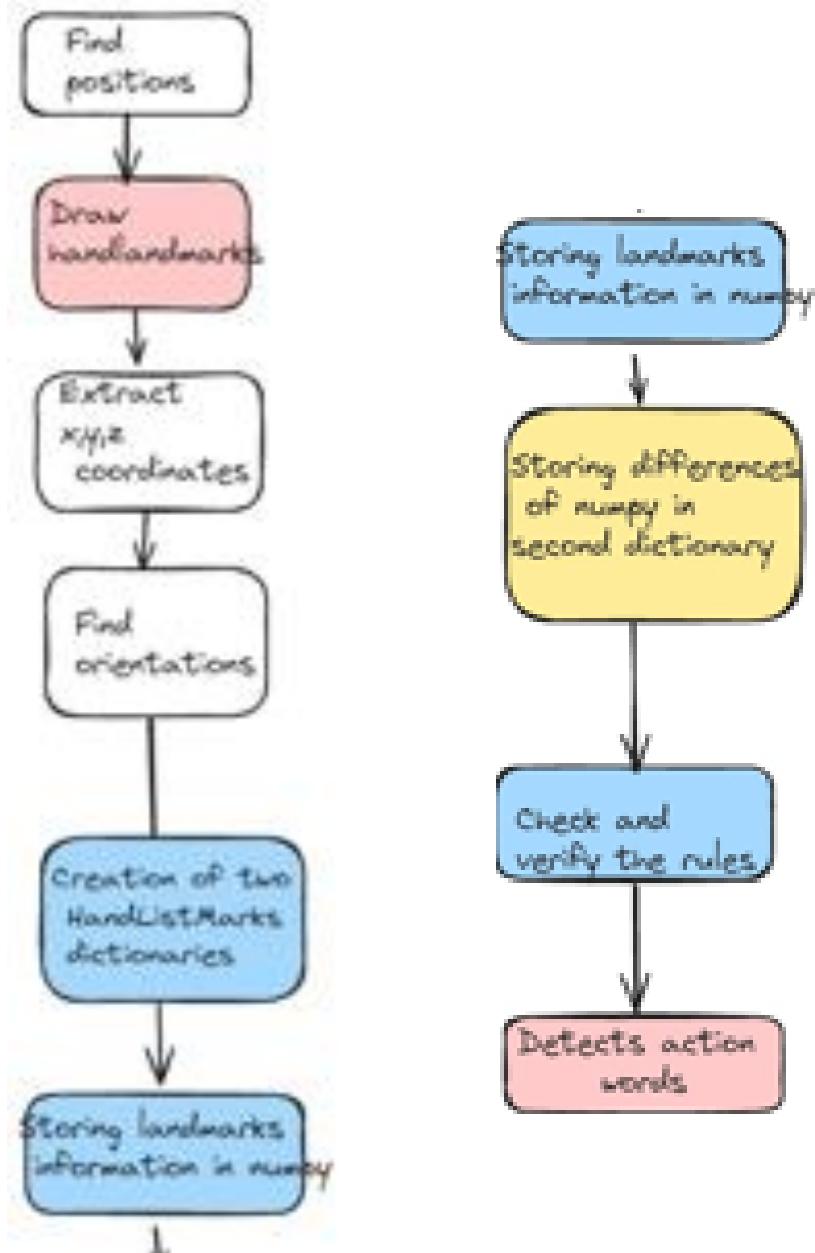
- Sequence of HandLandMarks are converted to Numpy arrays and stored in Python dictionaries.
- Differences of consecutive HandLandMarks also stored in Python dictionaries.
- By looking the x-coordinate values of differences we can determine in which X-direction the hand is moving or it is in still state with respect to X-direction
- Similar way the movements in three directions are determined to detect the action gestures

# Detecting ANSWER word by action

- The initial posture of ANSWER is the symbol of ‘A’
- The symbol of ‘A’ must be move forward towards the screen to show the word ANSWER. This can be determined by checking the differences in the Z-coordinate values.
- The movements in the X and Y coordinates are within the negligible values
- Similarly the symbol of ‘Q’ is moved forward to show the symbol for QUESTION.

# Detecting WHAT word by action

- The initial posture of WHAT is an open palm facing upward
- The palm is moved right to left at least two times is the action symbol for WHAT
- Here the changes in Y and Z directions are negligible.
- The movements in X-direction are significant
- By checking the differences in X-coordinates and if there are changes in the values from positive to negative and negative to positive at least two times that is the representation of the word WHAT





DIGITS



LETTERS



NUMBERS

ka— VIBRI

RELIANCE ENTERTAINMENT & PHANTOM FILMS PRESENT  
A KABIR KHAN FILMS PRODUCTION

NiNG RELIANCE

# INDIAN SIGN LANGUAGE

BY ISH NEWS

The central image is a movie poster for the film "83". It features a group of men in white cricket uniforms cheering and waving Indian flags. In the foreground, a woman in a pink top and a man in a grey shirt are smiling and making hand signs. The number "83" is prominently displayed in large, gold-colored digits. The "ISH NEWS" logo is at the bottom left. The poster includes text at the top and bottom.

PRODUCED BY DEEPIKA PADUKONE, KABIR KHAN, VISHNU VARDHAN INDURI, SAJID NADIADWALA,  
SHEETAL VINOD TALWAR, RELIANCE ENTERTAINMENT & 83 FILM LTD

1

2

3

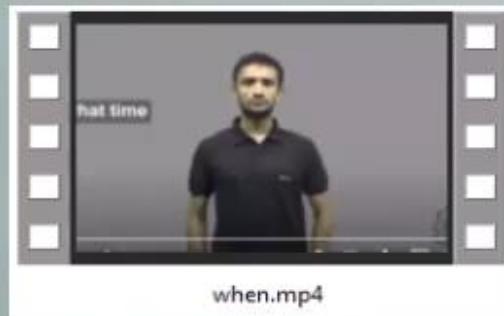
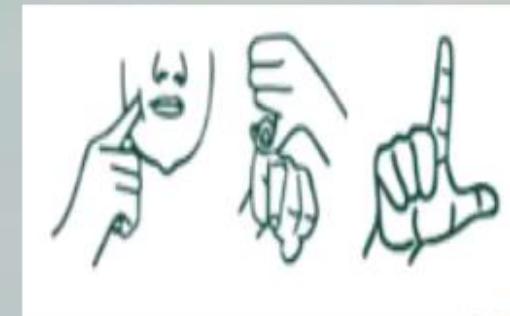
4

5

WORDS

SYMBOLS

EXIT

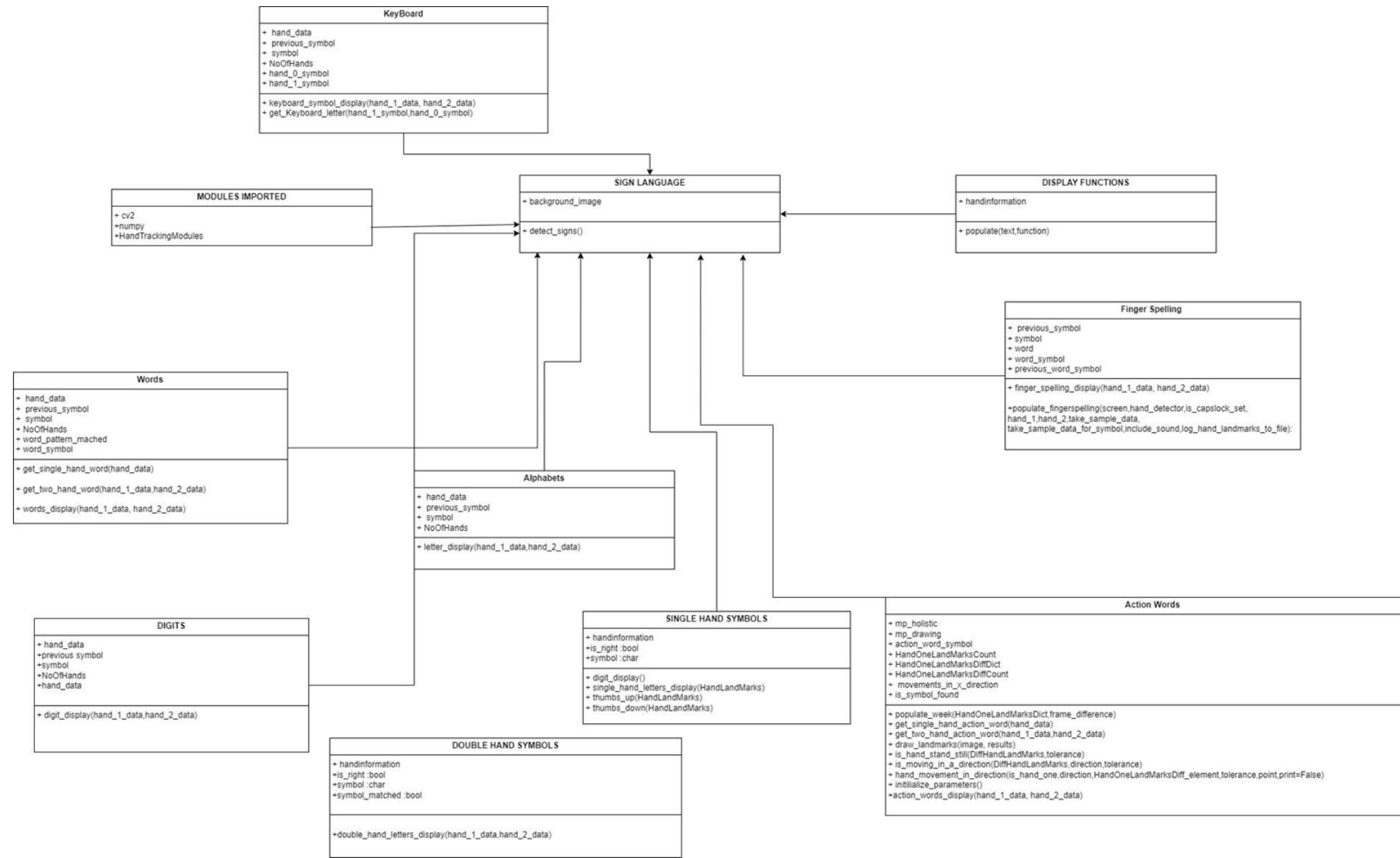


# REVIEW -5

Have implemented an extensible integrated working prototype of Real-Life Sign Language Detection

- NUMBERS
- ALPHABETS
- FINGER SPELLING
- STATIC SIGNS
- ACTION WORDS
- KEYBOARD IMPLEMENTATION USING SIGNS

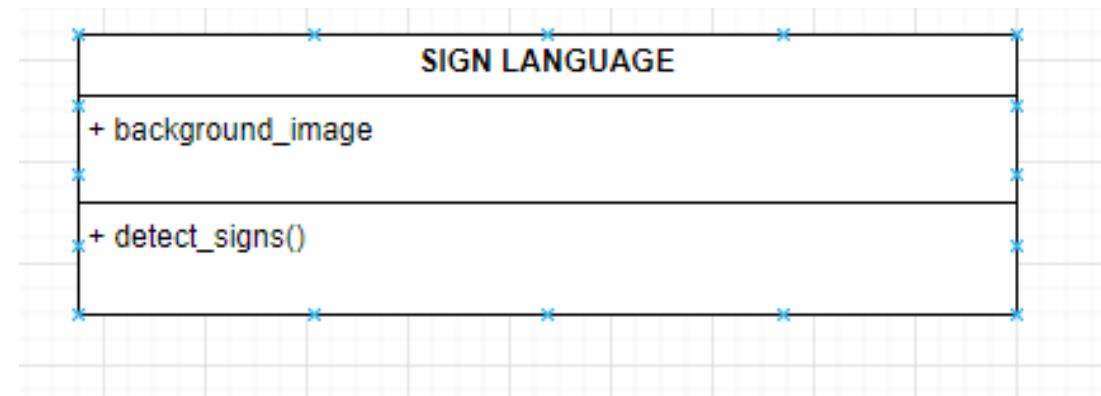
# IMPLEMENTATION DETAILS



# FIRST Component

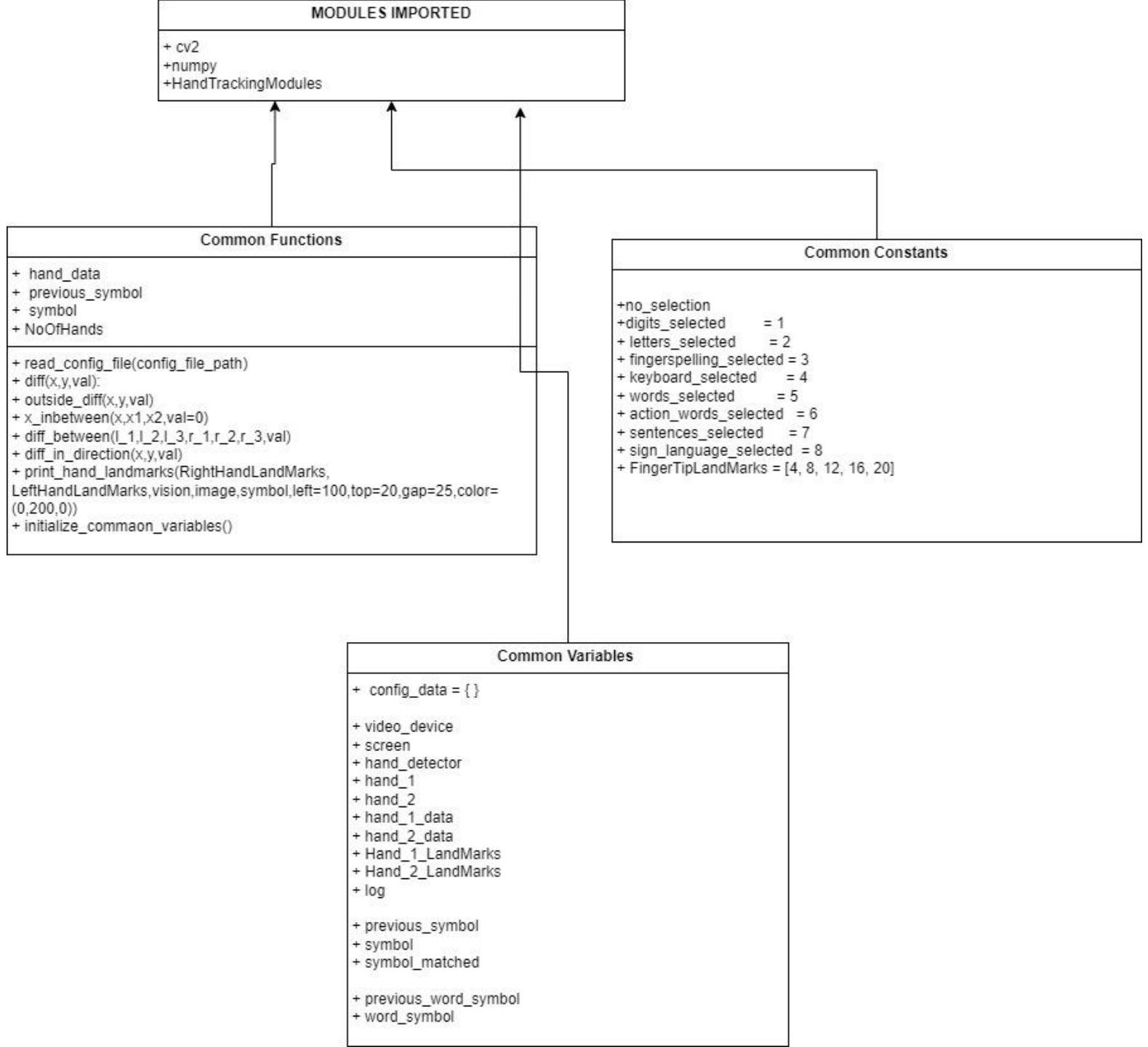
Has access to all modules

We decide where to go based on  
the module chosen



Module chosen is decided based on sign

# MODULES



# LOGGING

## LOGGING

- + hand\_1\_log\_count
- +hand\_2\_log\_count
- + log\_folder
- + log\_file\_name\_prefix
- + log\_file\_name\_suffix
- + log\_file\_path
- +hand\_one\_log\_file\_path
- + hand\_two\_log\_file\_path

- + create\_log\_files(self,symbol="")
- +open\_log\_file(self,symbol="")
- + write\_message(self,message,hand\_name="")
- + write\_hand\_landmarks(self,hand\_info,hand\_name="")

```
new 22 x double_hand_letters.py x new 24 x common_constants.py x common_functions.py x common_variables.py x new 25 x log_functions.py x Logging_what.txt x
1 Log started
2 First Right hand at 0 degrees
3 Thumb is Closed
4 Index finger is Closed
5 Middle finger is Closed
6 Ring finger is Opened
7 Little finger is Opened
8 HandLandMarks: [[0, 789, 507, 0], [1, 767, 509, -78], [2, 746, 508, -133], [3, 730, 511, -176], [4, 710, 515, -221], [5, 727, 494, -162], [6, 726, 459, -240],
9 Second Left hand at 270 degrees
10 Thumb is Opened
11 Index finger is Opened
12 Middle finger is Closed
13 Ring finger is Closed
14 Little finger is Closed
15 HandLandMarks: [[0, 796, 487, 0], [1, 834, 479, -128], [2, 860, 484, -209], [3, 870, 493, -271], [4, 886, 501, -334], [5, 839, 529, -214], [6, 826, 533, -319],
16 First Left hand at 0 degrees
17 Thumb is Opened
18 Index finger is Closed
19 Middle finger is Closed
20 Ring finger is Closed
21 Little finger is Closed
22 HandLandMarks: [[0, 770, 496, 0], [1, 759, 508, -101], [2, 743, 513, -162], [3, 731, 517, -205], [4, 711, 522, -249], [5, 725, 493, -181], [6, 719, 464, -250],
23 Second Left hand at 270 degrees
24 Thumb is Opened
25 Index finger is Closed
26 Middle finger is Closed
27 Ring finger is Closed
28 Little finger is Closed
29 HandLandMarks: [[0, 794, 490, 0], [1, 828, 482, -107], [2, 852, 487, -178], [3, 867, 495, -235], [4, 889, 503, -296], [5, 834, 527, -189], [6, 825, 534, -287],
30 First Left hand at 0 degrees
31 Thumb is Opened
32 Index finger is Opened
33 Middle finger is Opened
34 Ring finger is Opened
35 Little finger is Opened
```

Activate Windows

Go to Settings to activate Windows

```
double_hand_letters.py new 24 common_constants.py common_functions.py
1 [_
2 [0, 1060, 635, 0],
3 [1, 1001, 600, -64],
4 [2, 960, 549, -121],
5 [3, 930, 510, -185],
6 [4, 915, 465, -247],
7 [5, 1011, 450, -26],
8 [6, 978, 400, -155],
9 [7, 947, 412, -268],
10 [8, 927, 438, -328],
11 [9, 1054, 433, -58],
12 [10, 1050, 352, -136],
13 [11, 1035, 301, -206],
14 [12, 1018, 259, -251],
15 [13, 1096, 443, -110],
16 [14, 1113, 367, -184],
17 [15, 1119, 316, -224],
18 [16, 1120, 269, -247],
19 [17, 1131, 475, -170],
20 [18, 1156, 418, -220],
21 [19, 1173, 380, -220],
22 [20, 1187, 343, -215]
23 ]
```

# CONFIGURATION JSON

During Development there arised clashes in the already existent signs as well as the new signs which were being developed .

Ex: week and exit

Also there are important flags for logging and debugging which need access globally.

So also to give it a single point of access I made the configuration.json

So to eradicate this difficulty I made configuration.json

# CONFIGURATION JSON

In configuration.json you include all the existent signs and give them a value — “i” for “include”

— “e” for “exclude”

Only the sign names which are given value “i” will be able to participate in the sign detection for that duration and only their rules will be checked for sign prediction .Rest wont .

So this creates a temporary isolated environment suitable for developing rules of new signs and testing them .

We include Alphabets, Words, and action Words in this file

# CONFIGURATION JSON

Important Flags :

Logging

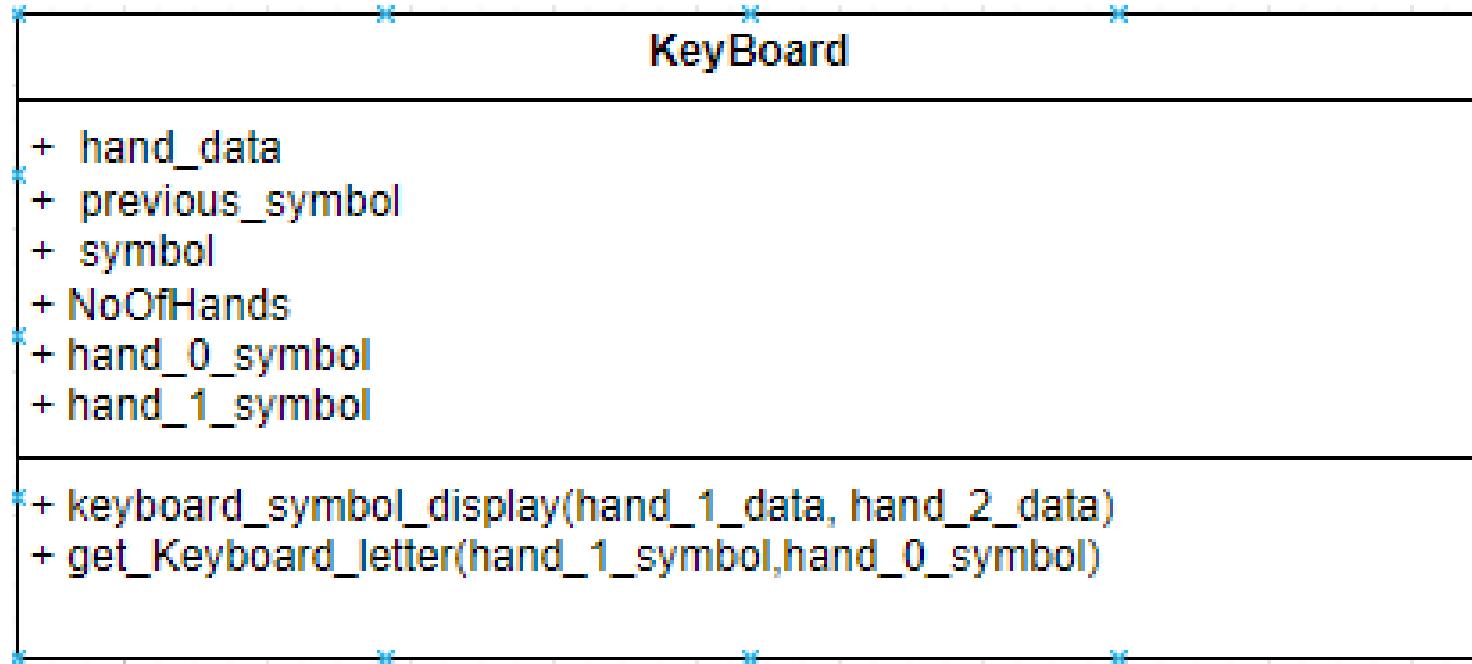
Debugging

Control flow

Testing

| Configuration                          |  |
|--|--|
| + alphabets : [ ]                      |  |
| + words : [ ]                          |  |
| + action words : [ ]                   |  |
| + webcam : bool                        |  |
| + debug : bool                         |  |
| + to_control_flow_for_debug : bool     |  |
| + default_option_for_debug : int       |  |
| + log_file_path : string               |  |
| + log_hand_landmarks_to_file : bool    |  |
| + take_sample_data : bool              |  |
| + take_sample_data_for_symbol : string |  |
| + include_sound : bool                 |  |

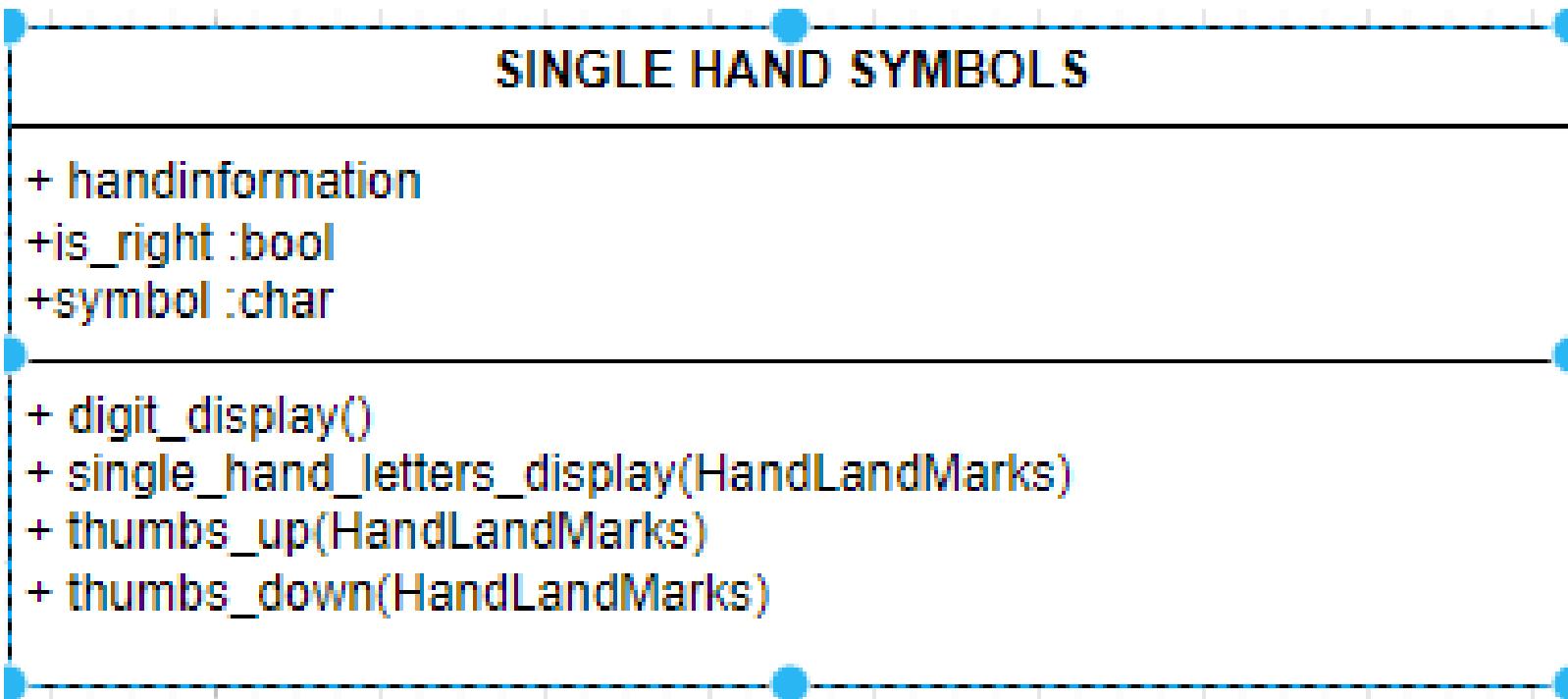
# KEYBOARD



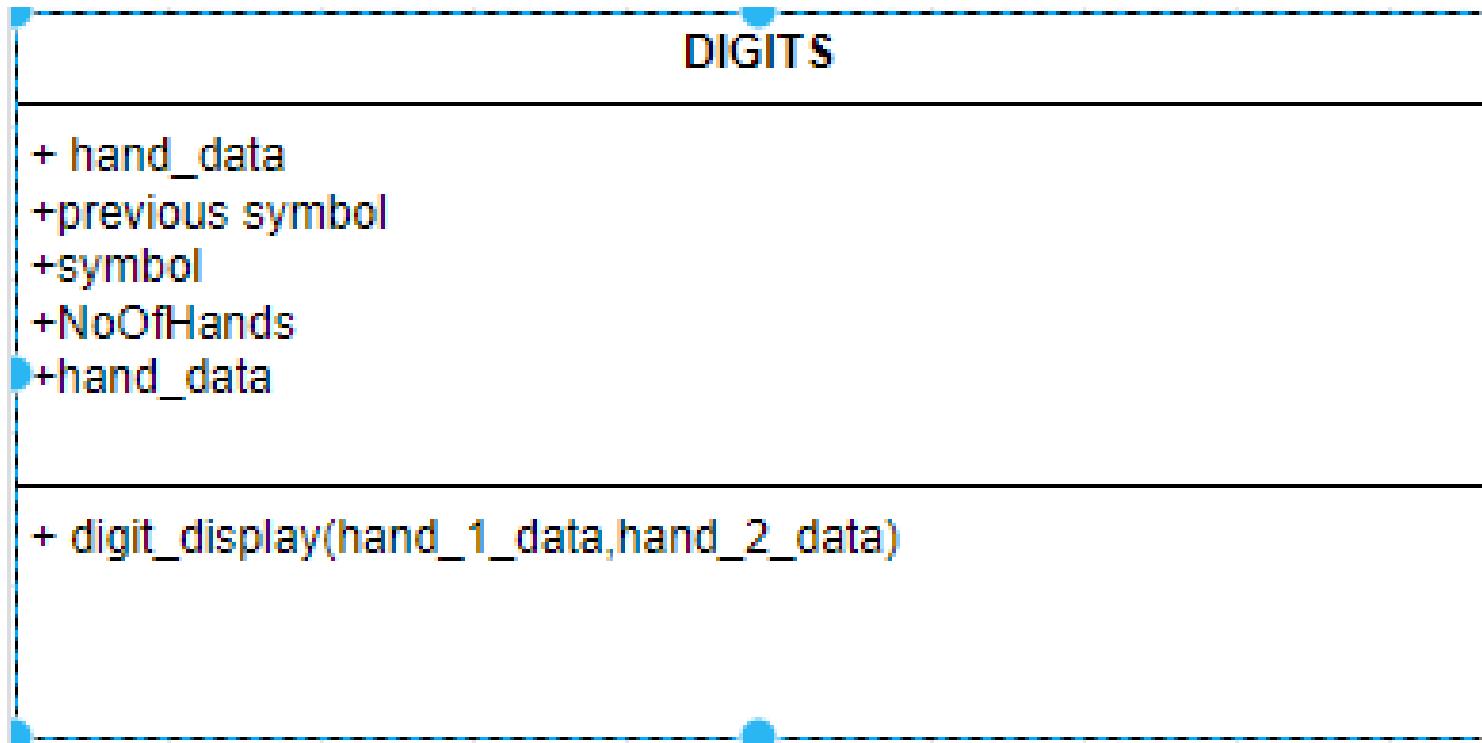
## Action Words



# SINGLE HAND SYMBOLS



# DIGITS



# ALPHABETS

## Alphabets

- + hand\_data
- + previous\_symbol
- + symbol
- + NoOfHands
  
- + letter\_display(hand\_1\_data,hand\_2\_data)

# DOUBLE HAND FUNCTIONS



---

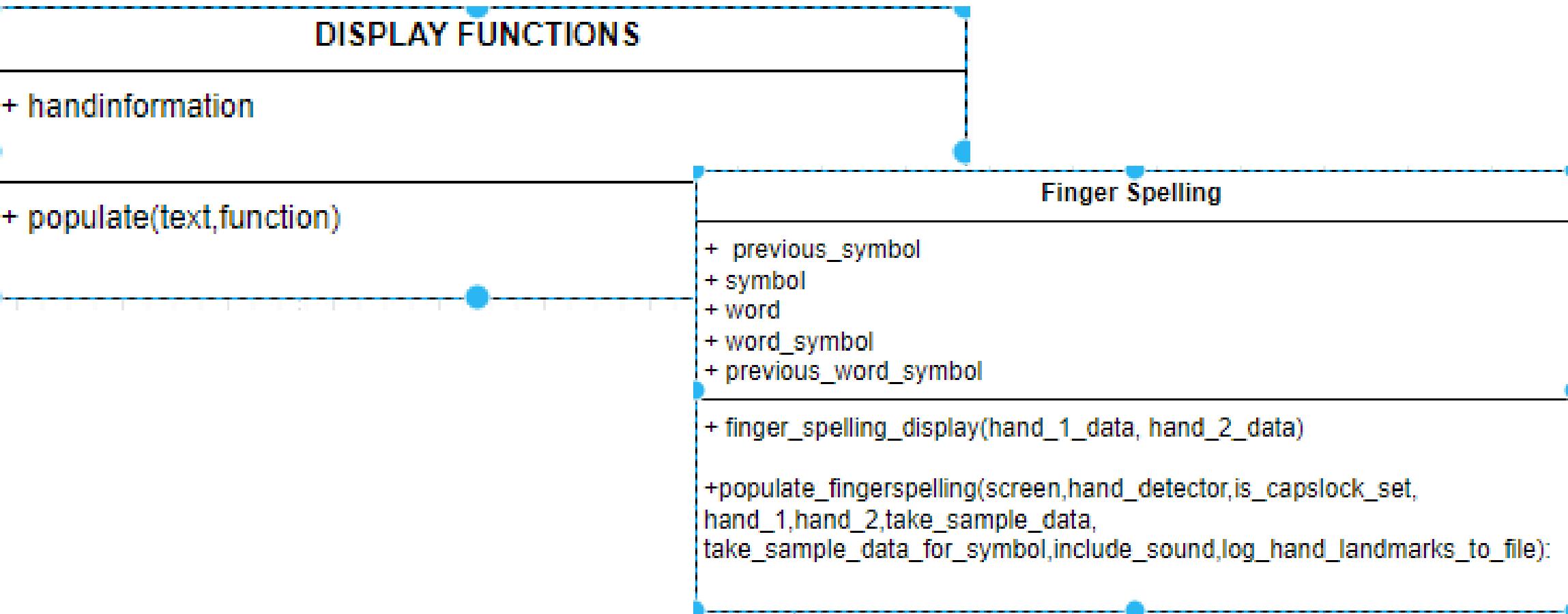
```
+ handinformation  
+is_right :bool  
+symbol :char  
+symbol_matched :bool
```

---

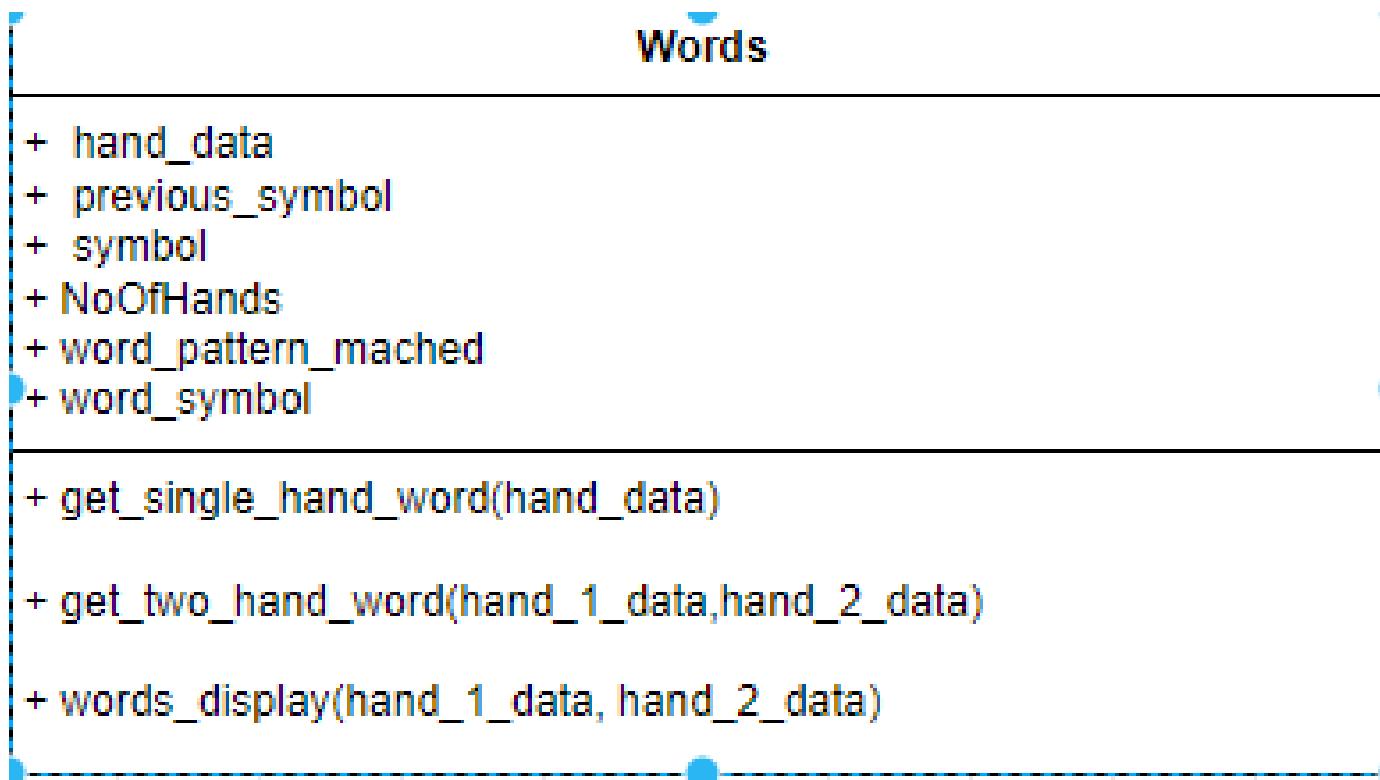
```
+double_hand_letters_display(hand_1_data,hand_2_data)
```



# DISPLAY FUNCTIONS and Finger Spelling



# WORDS



# ACTION WORDS

## Action Words

```
+ mp_holistic
+ mp_drawing
+ action_word_symbol
+ HandOneLandMarksCount
+ HandOneLandMarksDiffDict
+ HandOneLandMarksDiffCount
+ movements_in_x_direction
+ is_symbol_found

+ populate_week(HandOneLandMarksDict,frame_difference)
+ get_single_hand_action_word(hand_data)
+ get_two_hand_action_word(hand_1_data,hand_2_data)
+ draw_landmarks(image, results)
+ is_hand_stand_still(DiffHandLandMarks,tolerance)
+ is_moving_in_a_direction(DiffHandLandMarks,direction,tolerance)
+ hand_movement_in_direction(is_hand_one,direction,HandOneLandMarksDiff_element,tolerance,point,print=False)
+ initialize_parameters()
+action_words_display(hand_1_data, hand_2_data)
```

Active  
Go to S

## Action Words



# References

1. Anshul Mittal, Pradeep Kumar, Partha Pratim Roy, Raman Balasubramanian and Bidyut B. Chaudhuri. A Modified-LSTM Model for Continuous Sign Language Recognition using Leap motion. DOI 10.1109/JSEN.2019.2909837, IEEE Sensors Journal
2. Thad Starner, Student Member, IEEE, Joshua Weaver, and Alex Pentland, Member, IEEE Computer Society. Real-Time American Sign Language Recognition Using Desk and Wearable Computer Based Video. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 20, NO. 12, DECEMBER 1998
3. Muthu Mariappan H, Gomathi V. Indian Sign Language Recognition through Hybrid ConvNet-LSTM Networks. EMITTER International Journal of Engineering Technology Vol. 9, No. 1, June 2021, pp. 182~203 DOI : 10.24003/mitter.v9i1.613
4. B. NATARAJAN, E. RAJALAKSHMI, R. ELAKKIYA, KETAN KOTECHA, AJITH ABRAHAM, (Senior Member, IEEE), LUBNA ABDELKAREIM GABRALLA, AND V. SUBRAMANIYASWAMY. Development of an End-to-End Deep Learning Framework for Sign Language Recognition, Translation, and Video Generation. date of publication 28 September 2022, IEEE Access, date of current version 7 October 2022. Digital Object Identifier 10.1109/ACCESS.2022.3210543
5. CANDY OBDULIA SOSA-JIMÉNEZ 1 , HOMERO VLADIMIR RÍOS-FIGUEROA 2 , AND ANA LUISA SOLÍS-GONZÁLEZ-COSÍO 3. A Prototype for Mexican Sign Language Recognition and Synthesis in Support of a Primary Care Physician. IEEEAccess, Received 29 October 2022, accepted 17 November 2022, date of publication 5 December 2022, date of current version 9 December 2022. Digital Object Identifier 10.1109/ACCESS.2022.3226696

# References

6. MOHAMED A. BENCHERIF, MOHAMMED ALGABRI, MOHAMED A. MEKHTICHE, MOHAMMED FAISAL, MANSOUR ALSULAIMAN, HASSAN MATHKOUR, MUNEER AL-HAMMADI, (Member, IEEE), AND HAMID GHALEB. Arabic Sign Language Recognition System Using 2D Hands and Body Skeleton Data. IEEE Access, Received March 1, 2021, accepted March 14, 2021, date of publication March 30, 2021, date of current version April 26, 2021. Digital Object Identifier 10.1109/ACCESS.2021.3069714
7. Deep Kothadiya, Chintan Bhatt, Krenil Sapariya, Kevin Patel, Ana-Belén Gil-González and Juan M. Corchado. Deepsign: Sign Language Detection and Recognition Using Deep Learning. Electronics 2022, 11, 1780. <https://doi.org/10.3390/electronics11111780>
8. Divya Deora, Nikesh Bajaj. INDIAN SIGN LANGUAGE RECOGNITION. 2012 1st International Conference on Emerging Technology Trends in Electronics, Communication and Networking
9. QAZI MOHAMMAD AREEB, MARYAM, MOHAMMAD NADEEM, ROOBAEA ALROOBAEA, AND FAISAL ANWER. Helping Hearing-Impaired in Emergency Situations: A Deep Learning-Based Approach. IEEE Access, Received December 27, 2021, accepted January 8, 2022, date of publication January 13, 2022, date of current version January 24, 2022. Digital Object Identifier 10.1109/ACCESS.2022.3142918
10. Muthu Mariappan H and Dr Gomathi V. Real-Time Recognition of Indian Sign Language. Second International Conference on Computational Intelligence in Data Science (ICCIDDS-2019), 978-1-5386-9471-8/19/\$31.00 ©2019 IEEE

# References

11. E. RAJALAKSHMI<sup>1</sup>, R. ELAKKIYA, V. SUBRAMANIYASWAMY, L. PRIKHODKO ALEXEY, GRIF MIKHAIL, MAXIM BAKAEV, KETAN KOTECHA, LUBNA ABDELKAREIM GABRALLA, AND AJITH ABRAHAM (Senior Member, IEEE). Multi-Semantic Discriminative Feature Learning for Sign Gesture Recognition Using Hybrid Deep Neural Architecture. *IEEE Access*, Received 19 November 2022, accepted 12 December 2022, date of publication 2 January 2023, date of current version 9 January 2023. Digital Object Identifier 10.1109/ACCESS.2022.323367
12. Barathi Subramanian, Bekhzod Olimov, Shraddha M. Naik, Sangchul Kim, Kil-Houm Park & Jeonghong Kim. An integrated mediapipe-optimized GRU model for Indian sign language recognition. *Scientific Reports*, (2022) 12:11964, <https://doi.org/10.1038/s41598-022-15998-7>, nature portfolio.
13. Shiqi Wang, Kankan Wang, Tingping Yang, Yiming Li & Di Fan. Improved 3D-ResNet sign language recognition algorithm with enhanced hand features. *Scientific Reports*, (2022) 12:17812, <https://doi.org/10.1038/s41598-022-21636-z>, nature portfolio
14. ZHENXING ZHOU , (Graduate Student Member, IEEE), VINCENT W. L. TAM , (Senior Member, IEEE), AND EDMUND Y. LAM , (Fellow, IEEE). SignBERT: A BERT-Based Deep Learning Framework for Continuous Sign Language Recognition. *IEEE Access* Received October 23, 2021, accepted November 28, 2021, date of publication December 3, 2021, date of current version December 14, 2021. Digital Object Identifier 10.1109/ACCESS.2021.3132668
15. DEEP R. KOTHADIYA, CHINTAN M. BHATT, TANZILA SABA, (Senior Member, IEEE), AMJAD REHMAN, (Senior Member, IEEE), AND SAEED ALI BAHAJ. SIGNFORMER: DeepVision Transformer for Sign Language Recognition. *IEEE Access*, Received 23 October 2022, accepted 10 December 2022, date of publication 9 January 2023, date of current version 18 January 2023. Digital Object Identifier 10.1109/ACCESS.2022.3231130

# References

16. JESTIN JOY, KANNAN BALAKRISHNAN, AND SREERAJ M. SignQuiz: A Quiz Based Tool for Learning Fingerspelled Signs in Indian Sign Language Using ASLR. IEEE Access, Received December 28, 2018, accepted January 24, 2019, date of publication March 4, 2019, date of current version March 18, 2019. Digital Object Identifier 10.1109/ACCESS.2019.2901863
17. JIWEI HU, YUNFEI LIU , KIN-MAN LAM, AND PING LOU. STFE-Net: A Spatial-Temporal Feature Extraction Network for Continuous Sign Language Translation. IEEE Access, Received 16 December 2022, accepted 25 December 2022, date of publication 5 January 2023, date of current version 16 May 2023. Digital Object Identifier 10.1109/ACCESS.2023.3234743
18. Murat Taskiran, Mehmet Killioglu, and Nihan Kahraman. A Real-Time System For Recognition Of American Sign Language By Using Deep Learning. 978-1-5386-4695-3/18/\$31.00 ©2018 IEEE
19. Zhibo Wang, Senior Member, IEEE, Tengda Zhao, Jinxin Ma, Hongkai Chen, Kaixin Liu, Huajie Shao, Member, IEEE, Qian Wang, Senior Member, IEEE, Ju Ren, Member, IEEE. Hear Sign Language: A Real-time End-to-End Sign Language Recognition System. IEEE TRANSACTIONS ON MOBILE COMPUTING, 1536-1233 (c) 2020 IEEE
20. Jian Wu, Student Member, IEEE, Lu Sun, and Roozbeh Jafari, Senior Member, IEEE. A Wearable System for Recognizing American Sign Language in Real-Time Using IMU and Surface EMG Sensors. IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 20, NO. 5, SEPTEMBER 2016
21. Jeonghwan Cheon, Seungdae Baek and Se-Bum Paik. Invariance of object detection in untrained deep neural networks. Frontiers in Computational Neuroscience. 03 November 2022  
DOI 10.3389/fncom.2022.1030707

