iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** III **Month of publication:** March 2022

DOI: https://doi.org/10.22214/ijraset.2022.40923

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

Text to Speech and Language Conversion in Hindi and English Using CNN

Prof. Shweta Patil[1], Ms. Swapnali Dhande[2], Ms. Megha More[3], Ms. Sanika Dalvi[4]
[1]Professor, [2, 3, 4]Students Dept. of Computer Science, Shivajirao S. Jondhale college of Engineering, Dombivli, Maharashtra, India

Abstract: Deaf and dumb persons use sign language to communicate with other people in their society. Because sign language is the only means of communication for persons who are deaf or hard of hearing, it is mostly utilized by them. Ordinary folks are unfamiliar with this language. A real-time sign language recognition system has been developed in this article to allow those who do not know sign language to communicate with hearing-impaired people more readily. In this case, we employed American Sign Language. We have used American Sign Language in this paper. We introducing the development and implementation of an American Sign Language (ASL) derived from convolutional neural network. Deep Learning Method is used to train a classifier to recognize Sign Language and Convolutional Neural Network (CNN) is used to extract features from the images. We have also used Text-To-Speech Synthesis to convert the detected output into speech format. With use of MATLAB function the obtained text is converted into voice. In our system we are converting text to speech in Hindi language. Therefore hand gesture made by deaf and dumb people has been anatomized and restated into text and voice for better communication.
Keywords: Sign Language Recognition, Image Processing, Text-To-Speech, American Sign Language, Classification, Convolutional Neural Network, Deep Learning real-time

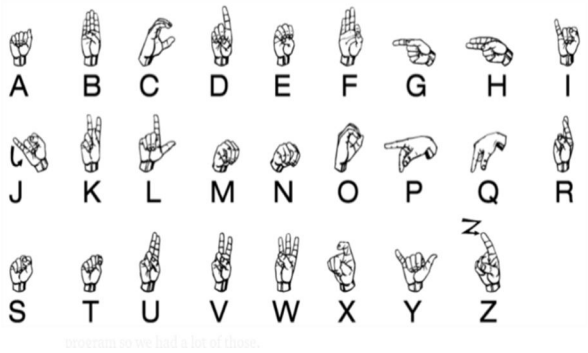
I. INTRODUCTION

As well said by Nelson Mandela, "Talk to a man in a language he understands, that goes to his head. Talk to him in his own language, that goes to his heart", language is undoubtedly essential to human interaction and has existed since human civilization began. Similarly sign language is a language used widely by people who are deaf-dumb, these are used as a medium for communication. Deaf and dumb people make use of their hands to expressing various gestures to expressing their thoughts and ideas with other people. Gestures are the nonverbally exchanged messages between people. Thus, this kind of nonverbal communication of deaf and dumb people is called **sign language.** A sign language is nothing but composed of colorful gestures formed by different shapes of hand, its movements. In different parts of the world different sign languages are used by people. Like American Sign Language (ASL), Indian Sign Language (ISL), Portuguese Sign Language (PSL), French Sign Language (FSL), British Sign Language (BSL) etc. We sometimes might not understand the language or the gestures the deaf-dumb people are trying to say.

Sign language being an essential thing for deaf-mute people, to interact both with normal people and with themselves, is still getting keen attention from the normal people. The importance of sign language is not being valued unless there are areas of concern with individuals who are deaf-mute. One of the answer to this situation is to talk with the deaf-mute people is by using the sign language. For non-verbal communication, hand gestures is used in wide range. It is mostly used by deaf & dumb people who have hearing or talking disorders to communicate among themselves or with other people. Over the years, many manufacturers have developed numerous sign language systems which are not supple and economical for customer. That is why end druggies can use sign language practitioner. But the usage of sign language interpreters could be expensive. Cost-effective result is needed so that the deaf-mute and normal people can communicate typically and easily. Our main aim is to implement an application which would detect predefined ASL via hand gestures. For the detection of movement of gesture, we would use basic level of hardware component like camera and interfacing is required. Our application would be a comprehensive User-friendly Based system built on PyQt5 module Instead of using technology like gloves or Kinect, we are trying to solve this problem using state of the art computer vision and machine literacy algorithms.

This application will comprise of two core module one is that simply detects the gesture and displays appropriate alphabet. The second is after a certain amount of interval period the scanned frame would be stored into buffer so that a string of character could be generated forming a meaningful word. Additionally, an add-on facility for the user would be available where a user can build their own custom-based gesture for a special character like period (.) or any delimiter so that a user could form a whole bunch of sentences enhancing this into paragraph and likewise. Whatever the predicted outcome was, it would be stored into a .txt file.

We have taken MNIST dataset from the following link on Kaggle https://www.kaggle.com/datamunge/sign-language-mnist.



II. LITERATURE REVIEW

The Purpose of the literature survey is to give the brief overview and also establish complete information about the reference papers.

1) *Volume: 08 Issue: 5,January 2020 ISSN : 2277-3878*
Translation Of Sign Language for Deaf a Dumb People

2) *Yildiz Technical University Istanbul, Turkey, Issue : 2018*
A Real-Time System For Recognition Of American Sign Language By Using Deep Learning

3) *Brandon Garcia Stanford University Stanford, CA*
Real-time American Sign Language Recognition with Convolutional Neural Networks.

4) *International Research Journal Of Engineering And Technology (IRJET) ,Volume: 9 Issue: 12, December 2020*
Sign Language Recognition System using Convolutional Neural Network and Computer Vision

5) *IOSR Journal of Computer Engineering (IOSR-JCE) ,Issue: 17, March 2018*
Speech to text and text to speech recognition systems

6) *Bangladesh University of Engineering and Technology Dhaka, Bangladesh, Issue: 31 October 2018*
Real-Time American Sign Language Recognition Using Skin Segmentation and Image Category Classification with CNN and Deep Learning

7) *International Journal of Research in Information Technology(IJRIT) ,Volume: 2 ,Issue: 5, May 2014*
Text – To – Speech Synthesis (TTS)

8) *Kennesaw State University Kennesaw, USA , Issue: 2018*
American Sign Language Recognition using Deep Learning and Computer

A. *Vision*

1) In this paper Image Processing and Artificial Intelligence are used, so we are taking this concept to develop algorithms. We use the hands to make different gestures that represent sign language which is captured as a series of images and MATLAB is used to process it and Output produces in the form of text
2) In this paper, a convolutional neural network is used as a fine classifier with TensorFlow and Keras libraries in Python. These Libraries work competently on significant modern GPUs (Graphics Processing Units) that permit doing much speedy computation and training.
3) From this research paper we have included the concept of CNN in our project . Basically they are used to translate a video of a user's ASL signs into text.
4) We are used the concept of Segmentation and resizing from this research paper. In this, the hand gesture is segmented firstly by taking out all the joined components in the image and secondly by letting only the part which is immensely connected, in our case is the hand gesture.

5) Machine Translation is a field of AI And NLP which uses machine translation system for converting different languages.
6) Image Classification-based on the output of the image features from previous step a classifier is trained using deep learning method to classify the input image with high accuracy. During testing phase, image that is saved is passed through feature.
7) Speech Application Programming Interface (SAPI) it is an interface between speech technology engines, both text to speech & speech recognition. It consists of 3 interface: Voice text interface, Attribute interface, Dialog Interface.
8) Gesture Detection- After creating the dataset of images for training the neural network to classify images, we can simply retain the existing inception model to work on our dataset.

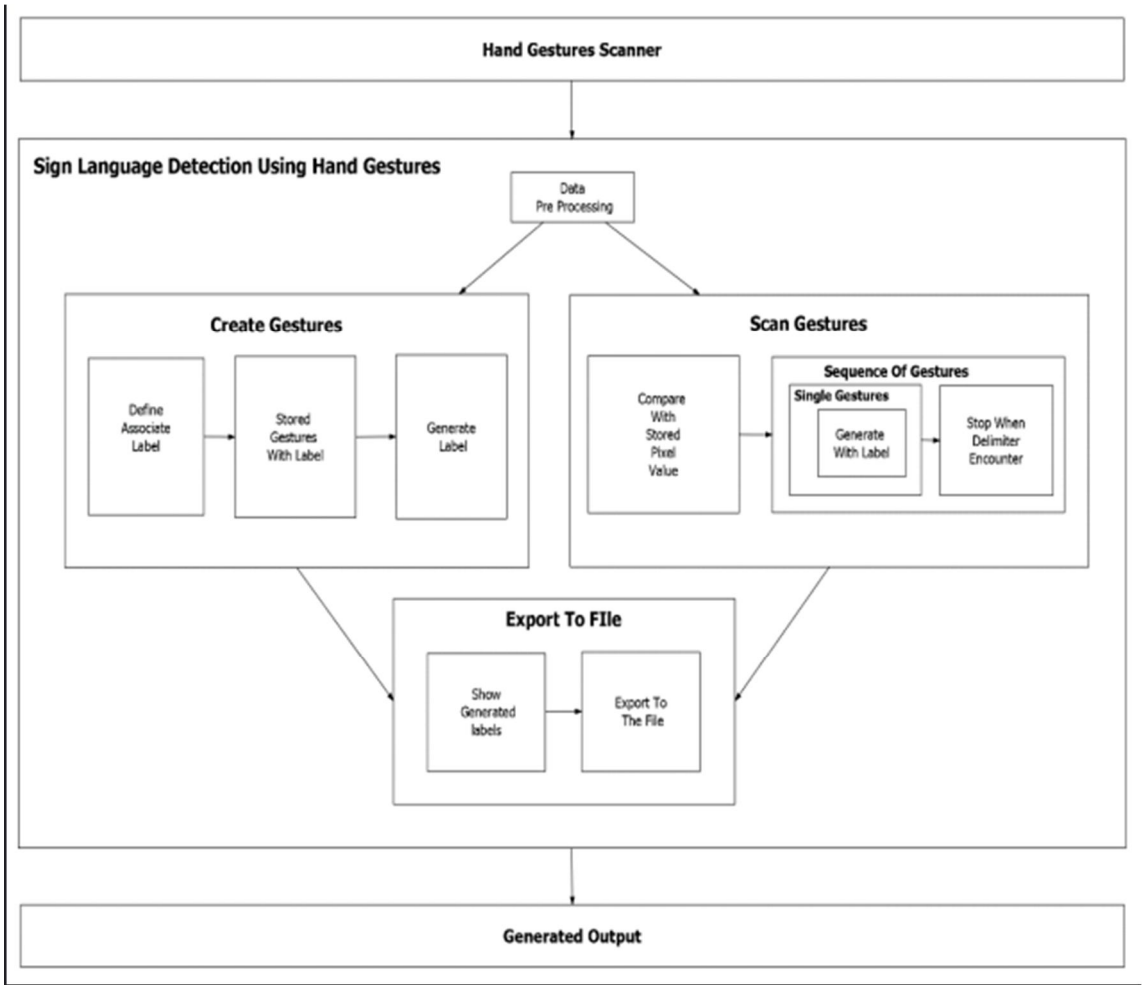
III. METHODOLOGY

A. *CNN*

In deep learning, a CNN is a class of deep neural networks, generally applied to dissect visual imagery.
In deep learning, a CNN is a class of artificial neural network, most commonly applied to analyze visual imagery.
We utilize convolutional neural network i.e. CNN for train the images and classify the images. Our model has achieved a remarkable accuracy of above 90% recognition sign language recognition, Hue Saturation Value (HSV) algorithm.

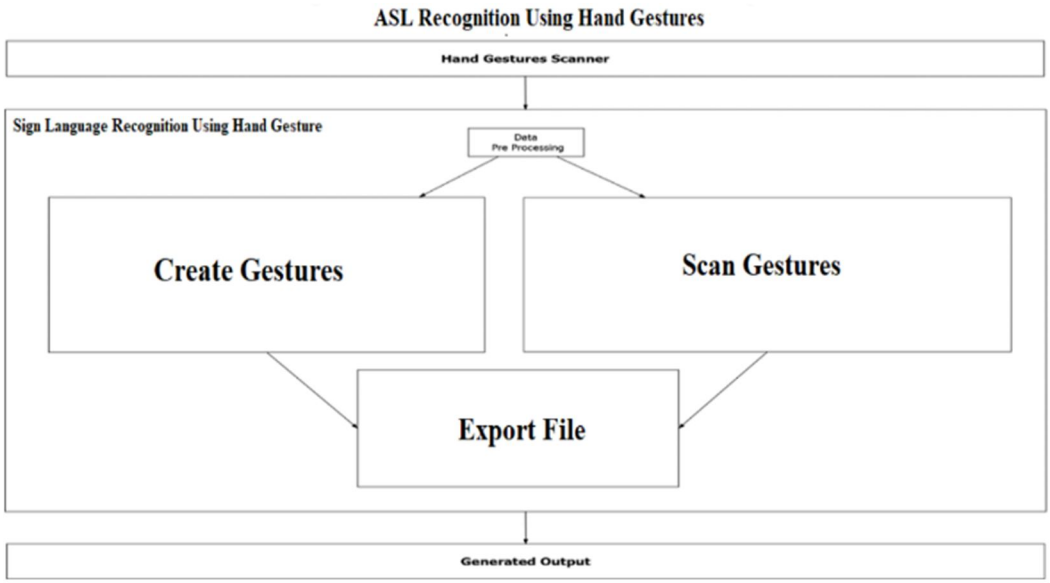
B. *System Architecture*

In our Project, there are two main modules. The frontend will be built on PyQT5 which will comprise of two core module one is that simply detects the gesture and displays appropriate alphabet. The second is after a certain amount of interval period the scanned frame would be stored into buffer so that a string of character could be generated forming a meaningful word. Additionally, we are trying to build our own custom-based gesture for a special character like period (.) or any delimiter so that a user could form a whole bunch of sentences enhancing this into paragraph and likewise. The predicted result would usually be stored into a .txt file



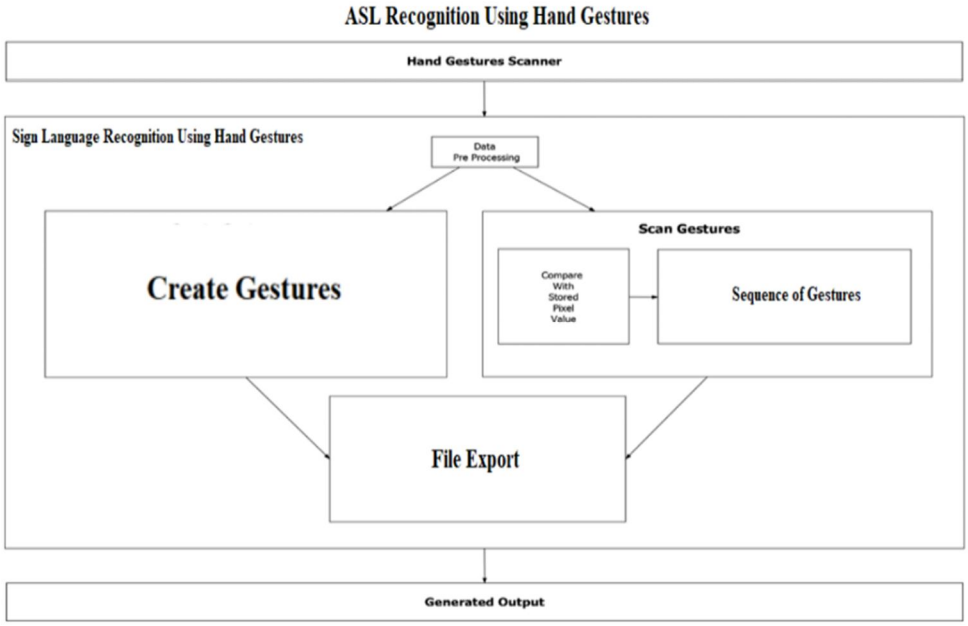
C. Modules in the System

1) Data Preprocessing



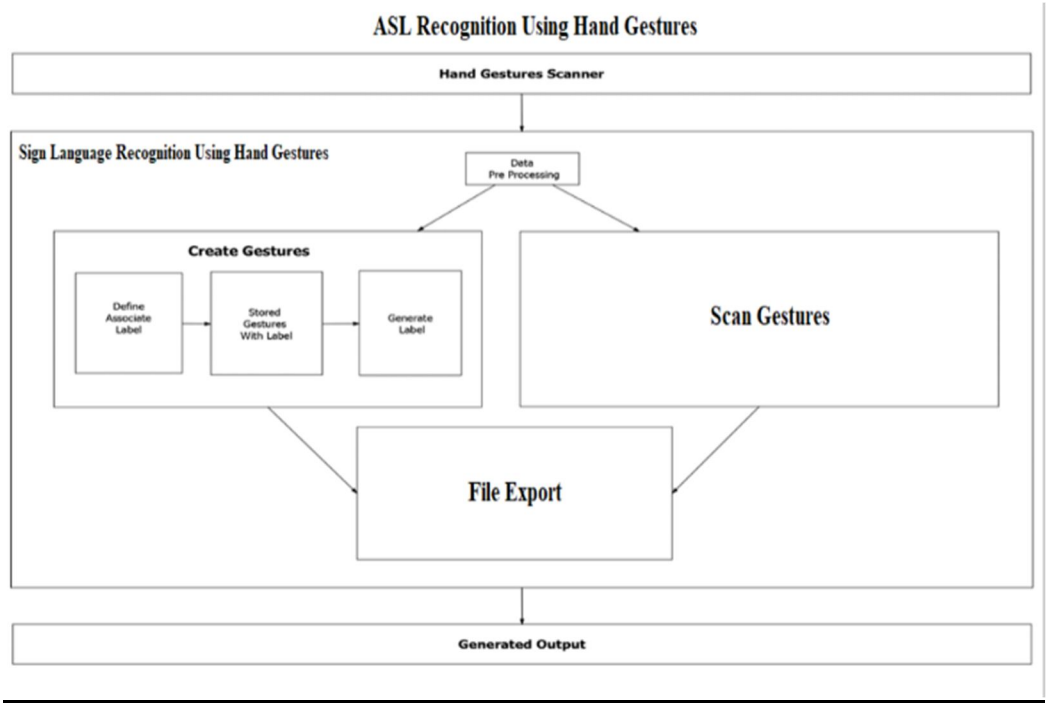
Preparing and processing the raw data to make it suitable for a ML model is called as Data Pre-processing. It is the first and most important step while creating a machine learning model. In this module, based on the object detected in front of the web cam it is converted into a binary image i.e. the object (hand) will be displayed in solid white and background will be filled with solid black. Based on the pixel's regions, their numerical value in range of either 0 or 1 is being given to next process for modules.

2) Scan Single Gesture



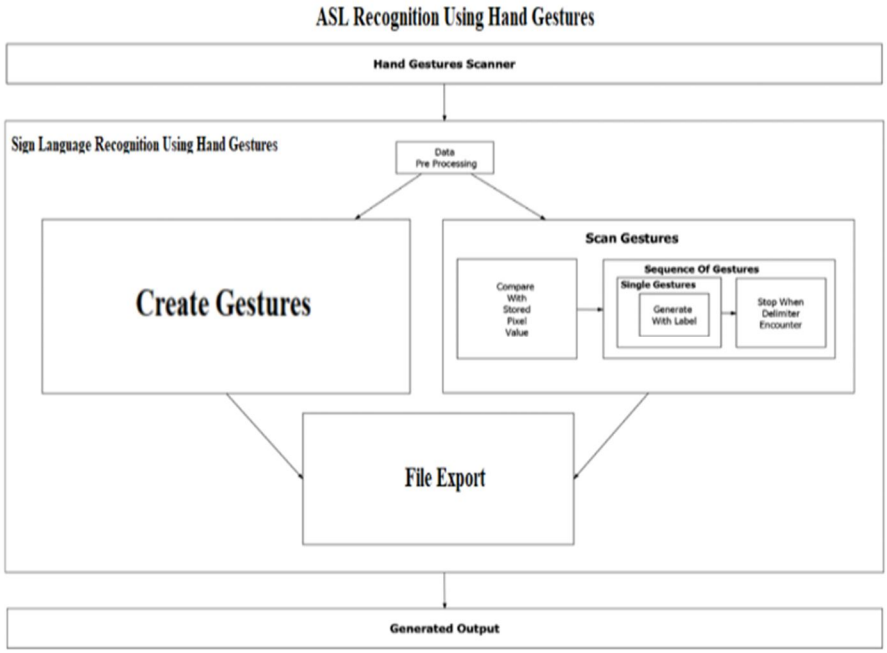
A gesture scanner will be available in front of the end user where the user will have to do a hand gesture. Grounded on Pre-Processed module output, user intended to see associated label allocate for each hand gestures, grounded on the predefined American Sign Language i.e. ASL standard inside the window of output screen.

3) Create Gesture



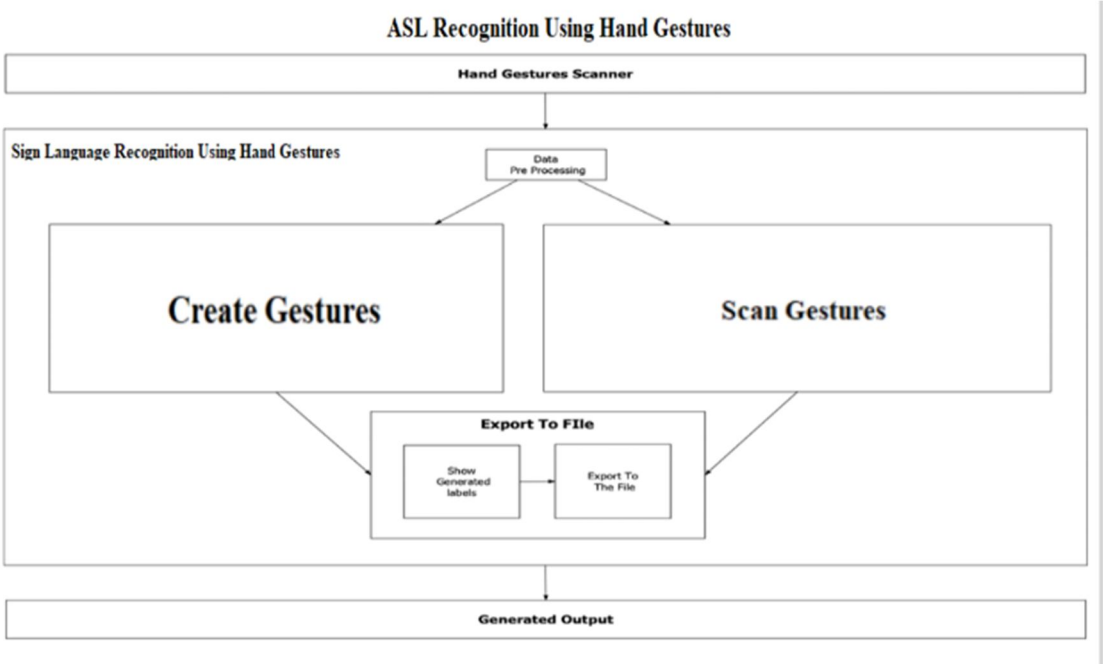
A user will give a desired hand gesture as an input to the system with the text box present at the bottom of the screen where the user needs to type whatever he/she wishes to label that gesture with. This customized gesture will then be stored for future purposes and can be detected when required.

4) Formation of a Sentence



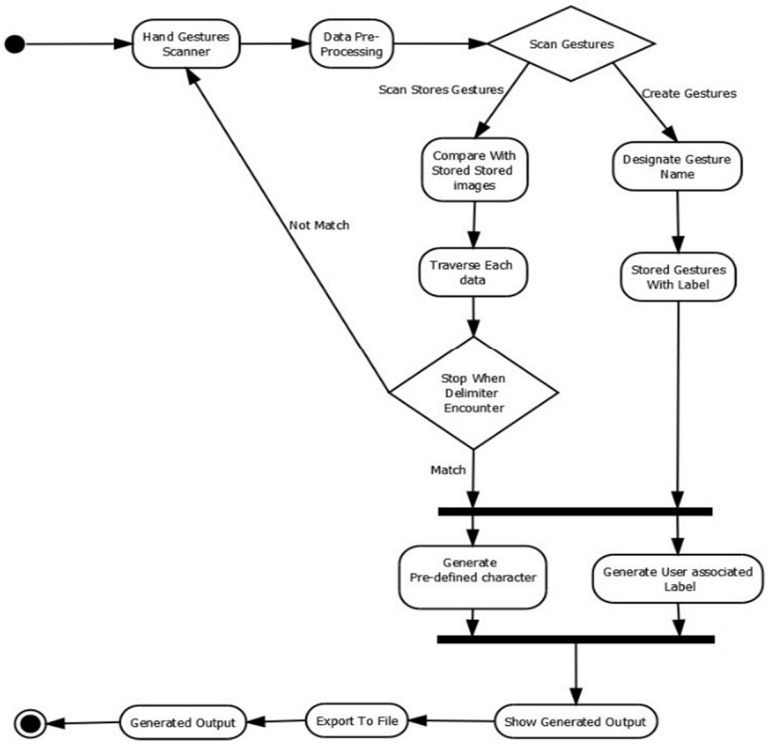
A delimiter will be selected by the user and until that delimiter is encountered every single scanned gesture will be appended with the previous result to form meaning-full words or sentences.

5) Exporting



A user would be able to export the results of the scanned character into an ASCII standard textual file format as well as into a speech.

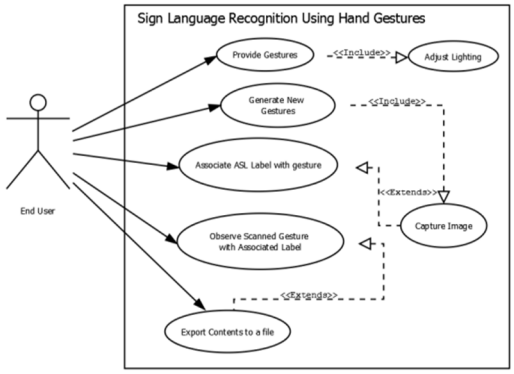
IV. DESIGN



Activity case diagram for sign language recognition using hand gesture

Firstly, we have to scan the hand gesture in scanner, after that data is going to be pre-processing then one decision has to be taken whether we have to scan the gesture or create the gesture. If we choose scanning the gesture, then our hand gesture will compare with stored dataset after that each data is travers and when delimiter encounter is stop then our gesture is matched with stored image then it will generate predefined character, if it is not matched then again we have to scan the gesture. In create gesture we have to create a new gesture then give a designated name to that gesture, after giving name it will stored that gesture with label. After we are done exporting the file, it will show the generated output, then we will be able to see the desired output in English as well as in Hindi language and also the text will be converted into speech.

A. Use Case Diagram

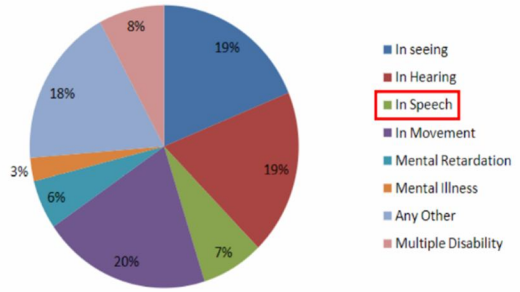


Use Case Diagram for Sign Language Recognition

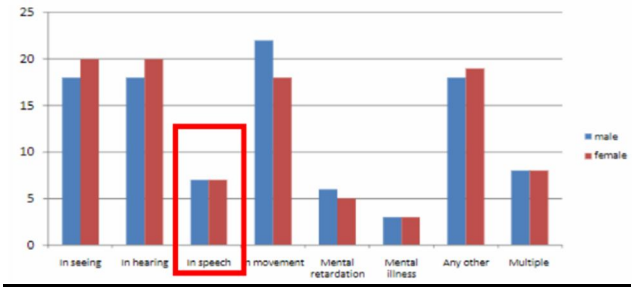
V. ANALYSIS

A. Statistics Retrieved from United Nation Statistic Division

B. Disable Population by Type Of Disability In India Census 2011



C. Distribution Disabled Person by Sex and Type of Disability (%) In India Census 2011



VI. CONCLUSION

From this application we have tried to overshadow some of the major problems faced by the disabled persons in terms of talking. We found out the root cause of why they can't express more freely. The result that we got was the other side of the audience are not able to interpret what these persons are trying to say or what is the message that they want to convey. Thereby this application serves the person who wants to learn and talk in sign languages. By help of these application individual can grasp various gestures and their meaning as per ASL standards. They can quickly learn what alphabet is assigned to which gesture. With formation of sentence, there is a provison of scanning single gesture. A user need not be a literate person if they know the action of the gesture, they can quickly form the gesture and appropriate assigned character will be shown onto the screen. Concerning to the implementation, we have used TensorFlow framework, with Keras API. And for the user feasibility complete front-end is designed using PyQT5. Appropriate user-friendly messages are prompted as per the user actions along with what gesture means which character window. With addition an export to file module is also provided with TTS(Text-To-Speech) cooperation meaning whatever the sentence was formed , user will be able to listen to it and he/she can export the sentence while watching what gesture was made by him/her during sentence formation. It is user friendly and has necessary options, which can be used by end user to carry out desired operation.

VII. FUTURE SCOPE

A. various search engines as well as texting applications could be used like Google, WhatsApp, etc. this would help illiterate people to chat to communicate with other people or search something on google with the help of gestures.

B. This project is working on image currently, further development can lead to detecting the motion of video sequence and assigning it to a meaningful sentence with TTS assistance.

C. It can also be implemented in applications such as Google Meet, Zoom, etc.

D. In government sites, there is no assistance for deaf and mute people for filling out forms, so our system can help with these issue.

REFERENCES

[1] PADMAVATHI. S, SAIPREETHY.M.S, V. Indian sign language character recognition using neural networks. IJCA Special Issue on Recent Trends in Pattern Recognition and Image Analysis, RTPRIA (2013).

[2] Mitra Sushmita and Acharya Tinku, "Gesture recognition: a survey", IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews, VOL. 37; NO. 3, May 2007

[3] Chih-Chung Chang and Chih-Jen Lin. 2011. "LIBSVM: A library for support vector machines", ACM Trans. Intell. Syst. Technol. 2, 3, Article 27 (May 2011), 27 pages

[4] Shobhit Agarwal, "What are some problems faced by deaf and dumb people while using todays common tech like phones and PCs", 2017 [Online]. Available: https://www.quora.com/What-are-some-problems-faced-by-deaf-and-dumb-people-whileusing-todays-common-tech-like-phones-and-PCs, [Accessed April 06, 2019].

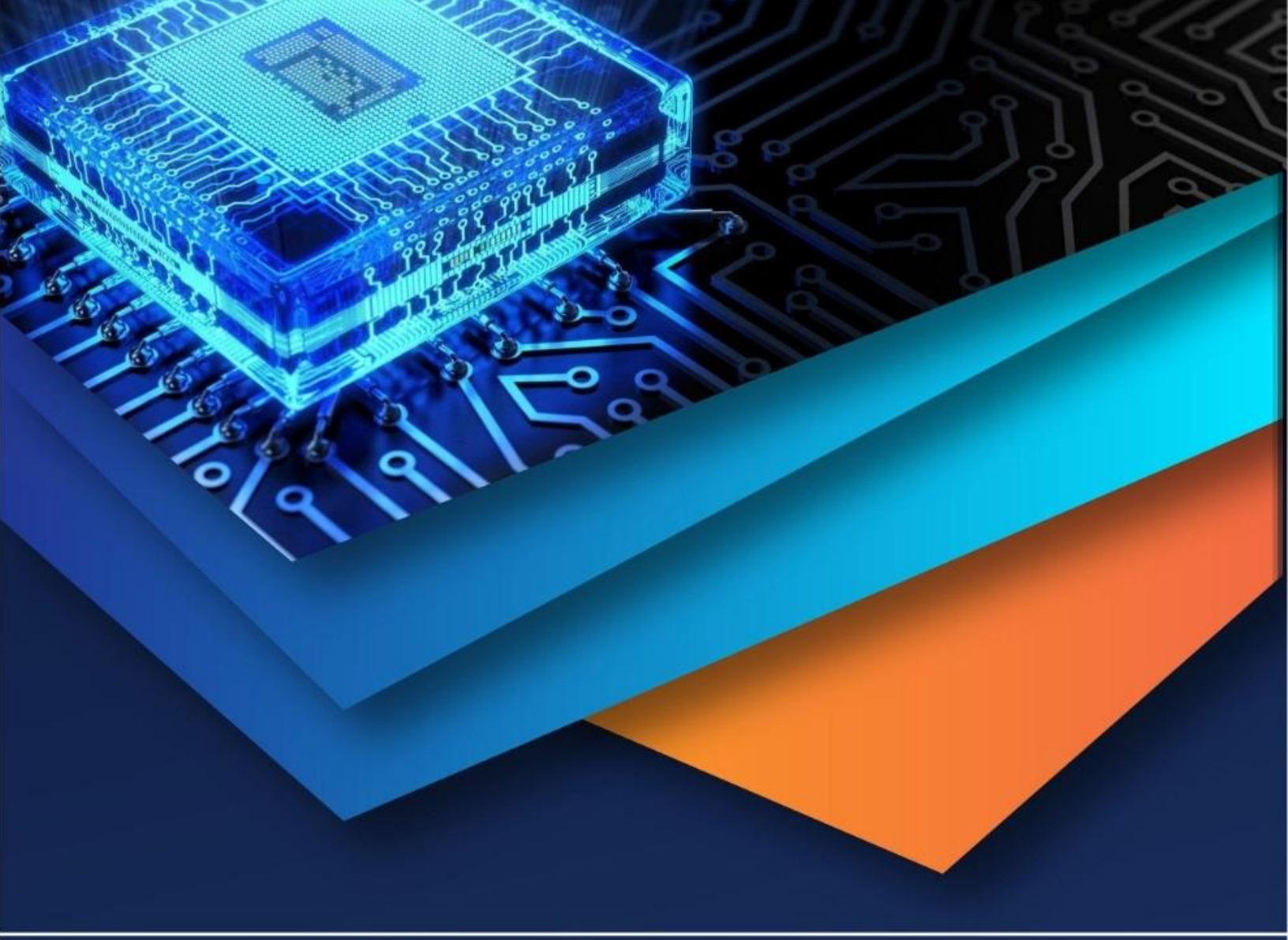
[5] M. Ibrahim, "Sign Language Translation via Image Processing", [Online]. Available: https://www.kics.edu.pk/project/startup/203 [Accessed April 06, 2019].

[6] NAD, "American sign language-community and culture frequently asked questions", 2017 [Online]. Available: https://www.nad.org/resources/american-sign-language/community-and-culturefrequently-asked-questions/ [Accessed April 06, 2019].

[7] Sanil Jain and K.V.Sameer Raja, "Indian Sign Language Character Recognition" , [Online]. Available: https://cse.iitk.ac.in/users/cs365/2015/_submissions/vinsam/report.pdf [Accessed April 06, 2019

[8] http://mospi.nic.in/sites/default/files/publication_reports/Disabled_persons_in_india_2016.pdf

[9] https://www.nidcd.nih.gov/health/american-sign-language



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)