**2.1 Introduction**

In India, around 10 million adults and 9 million children are affected by hearing loss and vision impairment. To aid deaf and blind children, we've developed a two-dimensional communication platform. Deaf or hard of hearing individuals have difficulty in communicating with others and expressing themselves. Deaf or partially deaf individuals have difficulty communicating with others and expressing themselves. They use hand gestures or sign language to convey themselves when speaking or communicating. Normal folks, on the other hand, find it difficult to comprehend these indications and gestures, resulting in a communication gap between the impaired and normal.

Sign Language is the primary mode of communication for the deaf and mute community. It arises and evolves naturally within hearing challenged communities. Two forms of sign are used for communication in sign language: manual and non-manual. A manual signs involve fingers, hands, arms and a non-manual signs involve face, head, eyes and body (Dixit, K., & Jalal, A. S., 2012).

Dixit, K., & Jalal, A. S. (2012). A Vision-Based Approach for Indian Sign Language Recognition. *International Journal of Computer Vision and Image Processing (IJCVIP)*, *2*(4), 25-36.

Sign Language is a well-structured language with a phonology, morphology, syntax and grammar (Waldron & Kim, 1995). (Waldron & Kim, 1995), It is a comprehensive natural language that uses diverse modes of expression for communication in everyday life. A Sign Language Recognition (SLR) system converts the conversation from human-human to human-computer interaction. These are used by deaf and dumb individuals to communicate with the hearing world. The purpose of the SLR system is to present an efficient and precise mechanism to transcribe text or speech, thus the “dialogue communication” between the deaf and hearing person will be seamless. There is no uniform sign language for all deaf individuals around the world. However, sign languages are not universal, as with spoken languages, these differ from region to region.

There are two basic ways utilised in the sign language recognition that is Glove/Device based and Vision based (Ong & Ranganath, 2005). (Ong & Ranganath, 2005). In the glove based method the user has to wear a gadget which bears a load of cables so as to connect the device to a computer. Such technologies are pricey and diminish the naturalness of the sign language conversation. In contrast, the Vision based technique uses simply a camera and directly deals with picture motions. It consists of two steps: sign capture and sign analysis. Vision-based solutions provide the user with a natural environment and decrease the complexities associated with glove-based methods. Extraction of visual data in the form of a feature vector is a crucial aspect of the gesture recognition challenge (Ong & Ranganath, 2005). However, there are obstacles such as hand tracking, hand separation from the environment and background, light fluctuation, occlusion, movement, and location (Stefan, Wang, & Athitsos, 2009).

High levels of grammatical variety exist in the sign languages of each country. Indian Sign Language is the prevalent name for India's sign language (ISL). It has been suggested that the same sign language is utilised throughout Nepal, Sri Lanka, Bangladesh, and Pakistan's border regions (Dasgupta, Shukla, Kumar, Diwakar, & Basu, 2008). Other sign languages include American Sign Language (ASL), British Sign Language (BSL), Korean Sign Language (KSL), etc. The All India Federation of the Deaf estimates that there are over 4 million deaf and over 10 million hard of hearing individuals in India (Dasgupta et al., 2008). One out of every five deaf people in the globe is from India, according to studies. More than one million deaf people and around half a million deaf children utilise Indian Sign Language as a form of communication; however, the majority of systems created for sign language identification are native language specific and cannot thus be used for ISL (Dasgupta et al., 2008).

In general, the semantic significance of language components in sign languages varies, but there are signals with universal grammar. A basic one-handed gesture conveying "hello" or "goodbye" has the same meaning around the world and in all sign languages. ISL is a natural language having its own morphology, phonology, syntax, and grammar. It is spoken in India (Dasgupta et al., 2008). ISL is a visual-spatial language that conveys linguistic information using hand, arm, facial, and head/body motions. ISL creates both discrete and continuous symptoms. An isolated sign concentrates on a single hand motion and consists of a precise hand configuration and stance depicted by a specific image. A continuous sign is a series of images representing a moving gesture.

**2.2 <Core area of the project> Could be clubbed with the Introduction**

The primary objective of the project is divided into two parts

1. In the first part the project will detect sign language using sequences in Real time and using multiple frames which will act as a sequence.

Will be done in 3 steps-

• Extracting holistic key points

• Training the neural network on the LSTM model

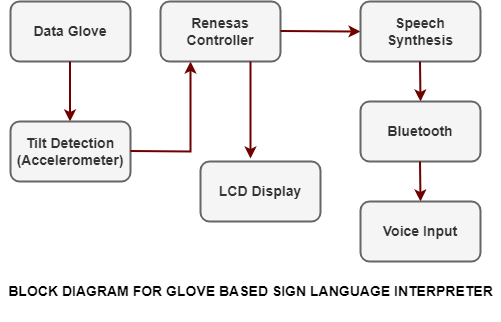
• Real time predictions using sequences

2. In the second part we propose an application which takes in live speech or audio recording as input, converts it into text and displays the relevant Indian Sign Language images or GIFs.

**2.3 Existing Approaches/Methods (Literature Review)**

2.3.1 Glove Based Method

All procedures in this project were carried out by referring to the Indian Sign Language (ISL) Portal. Fingerspelling is mostly employed in the ISL manual alphabet to spell out names or English terms that do not have an established sign. There are 26 ISL alphabets in the database, with the rest being static gestures. The performer's hand motions are captured using a low-cost hand glove circuit constructed with numerous accelerometers. It generates each finger's flexion, the hand's movement and orientation, and the electrical output from the hand's muscle actions. The system recognises gestures live, which means that the system takes the real-time signal from the gloves as an input and returns the matched gesture. It is entirely reliant on data.



**Overall Architecture of the Vision-based Sign Language Recognition System**

The Sign Language Recognition Prototype is a real-time vision-based system with the goal of recognising the Sign Language alphabet. The prototype's goal was to verify the validity of a vision-based system for sign language recognition while also testing and selecting hand features that could be employed with machine learning algorithms for usage in real-time sign language recognition systems.

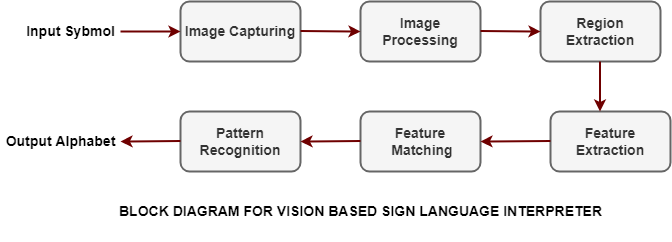
Trigueiros, P., Ribeiro, F., & Reis, L. P. (2014). Vision-based Portuguese sign language recognition system. In *New Perspectives in Information Systems and Technologies, Volume 1* (pp. 605-617). Springer, Cham.

To do so, the user must stand in front of the camera and make sign language movements, which the system will read and classify on the right side of the interface. Only one camera, a Kinect camera, is used in the implemented solution.

The entire work is carried out in two major stages. The fist stage consists of data gathering, preprocessing, and feature extraction; and the subsequent stage is classification of sign language gestures. The hand is recognised, tracked, and segmented from video pictures in the first module. Features for gesture categorization are derived from the segmented hand.

The resulting feature vector (instance vector) is normalised and classified in the gesture classification module using a previously trained Support Vector Machine (SVM), which is a pattern recognition technique in the field of supervised machine learning that works well with high-dimensional data (Bakheet, S., & Al-Hamadi, A., 2021).

Bakheet, S., & Al-Hamadi, A. (2021). Robust hand gesture recognition using multiple shape-oriented visual cues. *EURASIP Journal on Image and Video Processing*, *2021*(1), 1-18.

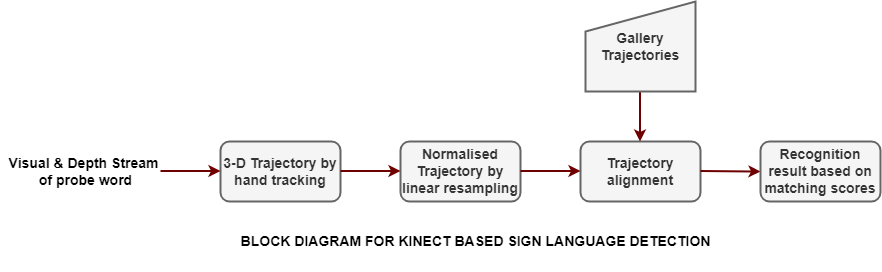


2.3.3   Kinect-Based Sign Language Hand Gesture Recognition System for Hearing- and Speech-Impaired

Individuals who are deaf or hard of hearing can use sign language to communicate with other members of society. Unfortunately, the majority of individuals do not understand sign language. A device based on image processing and pattern recognition could be useful in recognising and converting sign language into voice language in this case. This paper describes a system for using a custom-built software tool to identify and analyse sign language motions, and then translate the gesture into a voice language. The system uses a Dynamic Time Warping (DTW) algorithm to recognise specific gestures and an off-the-shelf software application to generate audible language.

Plouffe, G., & Cretu, A. M. (2015). Static and dynamic hand gesture recognition in depth data using dynamic time warping. *IEEE transactions on instrumentation and measurement*, *65*(2), 305-316.

Microsoft® Kinect is the primary instrument for capturing a user's video stream. With a 91 percent accuracy, the suggested technique can successfully detect gestures stored in the lexicon. Custom gestures can be defined and added to the proposed system. According to an experiment in which ten persons with disabilities used the system to communicate with five others who did not have a handicap, 87 percent believed that it was useful.



Name the Flow Chart

**2.4 <Pros and cons of the stated  Approaches/Methods >**

Glove based Approach

Glove based system supports a wider range of signs and modes. **Also,** this method uses a wireless transceiver system which was developed using a Gesture Vocalizer.

Havalagi, P. S., & Nivedita, S. U. (2013). The amazing digital gloves that give voice to the voiceless. *International Journal of Advances in Engineering & Technology*, *6*(1), 471.

Vision-based Sign Detection

The Vision-based gesture extraction method is accurate for face and body gestures; however**,** it has a few issues. It is often associated with image noises from various sources such as light, camera, colour matching and the background. Much prior research has been done in which the error filter failed to reconstruct the damaged parts due to its dynamic environment. Furthermore, this method also requires a vast computation along with a real time vision system.

Sarkar, A. R., Sanyal, G., & Majumder, S. J. I. J. O. C. A. (2013). Hand gesture recognition systems: a survey. *International Journal of Computer Applications*, *71*(15).

Kinect-Based (Microsoft Kinect XBOX 360TM)

Kinect based method was developed initially for gaming purposes. But later Microsoft Kinect became widely recognised due to numerous **industrial applications** in the fields of Computer Vision such as Gesture and Action recognition, VR(Virtual Reality), and Robotics. Although, this method is less widely used in industry due to its delayed responses when **the** feature size of image is large. **Otherwise,** it is quite an accurate method.

Han, J., Shao, L., Xu, D., & Shotton, J. (2013). Enhanced computer vision with microsoft kinect sensor: A review. *IEEE transactions on cybernetics*, *43*(5), 1318-1334.

**2.5 Issues/observations from investigation**

Glove based, Vision-based, and Kinect-Based methods either provide lesser accuracy or are more costly and **time-consuming** as compared to the sign language detection using LSTM. Moreover, LSTM provided a 90% accuracy using the first architecture. On the other hand, the glove based method provided only 75% accuracy. Furthermore, both Vision based and Microsoft Kinect methods are more expensive as compared to LSTM gesture detection.

Franslin, N. M. F., & Ng, G. W. (2022). Vision-Based Dynamic Hand Gesture Recognition Techniques and Applications: A Review. In *Proceedings of the 8th International Conference on Computational Science and Technology* (pp. 125-138). Springer, Singapore.

**2.6 Summary**

In this paper, we gave an overview of sign language recognition using neural networks and HMM techniques. The most common methods for collecting hand motions in sign language are image-based and hand glove-based systems combined with a hand tracker. Using data-gloves allows for more complex motions, including individual fingers, wrists, and hands, resulting in more flexible, accurate, and reliable gesture recognition. The image-based method, on the other hand, allows for user-independent feature extraction while requiring more processing for feature extraction and noise reduction.

Some studies that deal with sign language recognition using various methodologies (Neural Networks and HMM) systems are examined in length, with their benefits and drawbacks highlighted. Different Neural Network systems are employed in different stages, hence a comparison was done between each of these methods. A variety of recognition methods are based on the nature of the situation, complexity, as well as available surroundings. Most well-known 199 | 2013 International Conference on Computer Science and Information Technology Proceedings (CSIT-2013). The information presented is based on the author's personal data collecting as well as a limited vocabulary, the findings will not accurately reflect the situation. The system's dependability appears to be intriguing to set rather than using performance measures in sign recognition systems. The accuracy of words was checked in this report.

Alalyani, N., & Marie-Sainte, S. L. (2018). NADA: New Arabic dataset for text classification. *International Journal of Advanced Computer Science and Applications*, *9*(9).

Furthermore, because authentic human gestures are continuous, establishing a separate system can greatly interrupt the natural flow of human contact and thus has less utility in the real world of sign identification. A completely automated sign recognition system's success hinges on resolving current issues with continuous gesture recognition.

Waldherr, S., Romero, R., & Thrun, S. (2000). A gesture based interface for human-robot interaction. *Autonomous Robots*, *9*(2), 151-173.

**Literature review**

**Sign Language Recognition Using CNN and LSTM**

According to the 2011 census, around 63 million people suffer from hearing and listening issues, and they are treated as insignificant by the general public. It is critical to raise public knowledge about sign language, and one approach to doing so is to encourage sign language education among children at all levels of education, including elementary, secondary, and higher education. Many academics have already done extensive work on this subject for many sign languages, including American Sign Language (ASL), Italian gestures, Chinese Sign Language (CSL), and Arabic Sign Language (ArSL).

Chong, T. W., & Lee, B. G. (2018). American sign language recognition using leap motion controller with machine learning approach. *Sensors*, *18*(10), 3554.

As a result, a standardised dataset of sign languages is lacking. Previous work on sign language was primarily based on reputable studies, which classified hand detection algorithms into two categories: appearance-based and model-based. Fingertips were detected using an appearance-based method to enable hand gesture detection. This method uses a neural network-based system to recognise continuous hand positions from grey-scale video images. On the other hand, El-Sawah et al. used a histogram to calculate the chance of skin colour observation in the model-based approach. Hand identification has been proposed using Artificial Neural Networks (ANNs)/ learning-based approaches, fuzzy logic, and genetic algorithm-based strategies.

Kukker, A., & Sharma, R. (2021). A genetic algorithm assisted fuzzy Q-learning epileptic seizure classifier. *Computers & Electrical Engineering*, *92*, 107154.

 Dardas and Georganas created a hand gesture detection and identification system using the BOF approach and a multiclass SVM classifier. They developed a syntax for generating gesture commands that may be used to control apps. Their system may provide adequate real-time performance as well as excellent classification accuracy under changing scenarios. Their method, on the other hand, can only detect and track static postures. Their immobile postures prevented their grammar from forming sentences. Furthermore, the system's adaptability may be limited by appearance-based approaches, depth of field, and hand posture information.

The proposed analysis began with the creation of a corpus for the system to intelligently forecast a sign. For this goal, the author experimented with several datasets to identify properties such as matching points, edges, nodes, and gesture movement. On the other hand, the lack of abundant characters and words from the Indian dictionary was a big roadblock to machine training. Second, while some of the datasets were available with all of the necessary and sufficient terms, they were collected for purposes that were utterly unrelated to the author's objectives, namely, communication between the disabled and the general public. Finally, assuming the given datasets were overcoming the aforementioned challenges, the most serious of all issues would arise: the image quality was insufficient to feed into a CNN-LSTM network; yet, the author created this specific method for the following reasons:

 (i) Using the gesture's feature matrix as the corpus would have resulted in a 5-fold delay because the feature matrix must be turned into an image and vice versa.

Hernández-Vela, A. (2015). From pixels to gestures: learning visual representations for human analysis in color and depth data sequences.

(ii) The CNN-LSTM approach is proven to be the ideal algorithm for processing images for machine learning, as classic CNN is required and sufficient for a single image, while LSTM can store the memory of the last processed corpus and is suitable for numerous images.

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. *ieee Computational intelligenCe magazine*, *13*(3), 55-75.

**Speech to Sign Language translator**

Hearing-impaired people will find it challenging to communicate with normal people. The Speech to Sign Language translator is used to bridge the communication gap between normal and hearing-impaired people. This translator makes it easier and faster for normal people to communicate their ideas to people who are deaf or hard of hearing. Natural Language Processing techniques are used by the translator to transform speech or text into Sign Language. As a result, the system is being used to overcome the barriers that normal individuals encounter in sharing their thoughts with hearing-impaired people, and it will serve as an ear for the deaf.

The existing methods concentrate on word-by-word translation and translation to American Sign Language (ASL). It is difficult for Indian deaf people to comprehend sign languages from other nations, such as American Sign Language (ASL), British Sign Language (BSL), New Zealand Sign Language (NZSL), and others.

Our proposed solution is intended to alleviate the difficulties that Indian deaf people experience. The goal of this system is to convert each word received as input into sign language. This project uses Indian Sign Language to translate the words.

 • Natural Language Processing-Filler words in sign language conversion, such as "is," "are," "was," "were," etc., scarcely contribute to the context. As a result, the filler words are removed from the speech or sentence by the algorithm.

Birdsong, D. (2006). Dominance, proficiency, and second language grammatical processing. *Applied Psycholinguistics*, *27*(1), 46.

• Root Words: The words can be gerunds, plurals, or adjectives. The proposed approach will eliminate these forms of the words and look for the root term. These root words will aid in the transfer of speech to sign language.

Mikheev, A. (1997). Automatic rule induction for unknown-word guessing. *Computational Linguistics*, *23*(3), 405-423.

• Database: The system includes a large database of Indian sign language words that can be converted to text or speech. As a result, it will benefit all deaf individuals in India. It allows people to comprehend the majority of the speech or material.

Badhe, P. C., & Kulkarni, V. (2015, November). Indian sign language translator using gesture recognition algorithm. In *2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS)* (pp. 195-200). IEEE.

**Speech Recognition (Audio/speech to Sign language)**

The microphone in our system receives the live voice as input. To accomplish this, PyAudio, a Python package, is used. PyAudio is a Python tool for recording audio across several platforms. The Google Speech Recognizer API is used to transform the audio into text. It's an API that incorporates neural network models to help translate audio to text. The received audio is translated into text using this Google Speech Recognizer in the input format of supplying the audio file. The audio is broken into smaller portions based on the occurrence of quiet in longer audio files. The pieces are then fed through the Google Speech Recognizer, which converts them into text quickly.

**3.1 Introduction**

Object Detection is a computer technology that deals with image processing and computer vision, it detects and identifies objects of various types such as humans, animals, fruits & vegetables, vehicles, buildings etc. Every object in existence has its own unique characteristics which make them unique and different from other objects. Similarly we have tried to apply the LSTM to capture sign language and detect words out of it. The basic requirements are our own dataset and technical requirements.

**3.3 Specific Project requirements**

*3.3.1 Data requirement*

For the Module Signify the data is stored in a folder. Creation of different signs is done for the database as a comparison. 30 videos for each word were taken to compare and render the difference. The code for database creation was then run to create the sample image in the database.

Each image was stored as different folders to compare with different words in the main code after feature extractions.

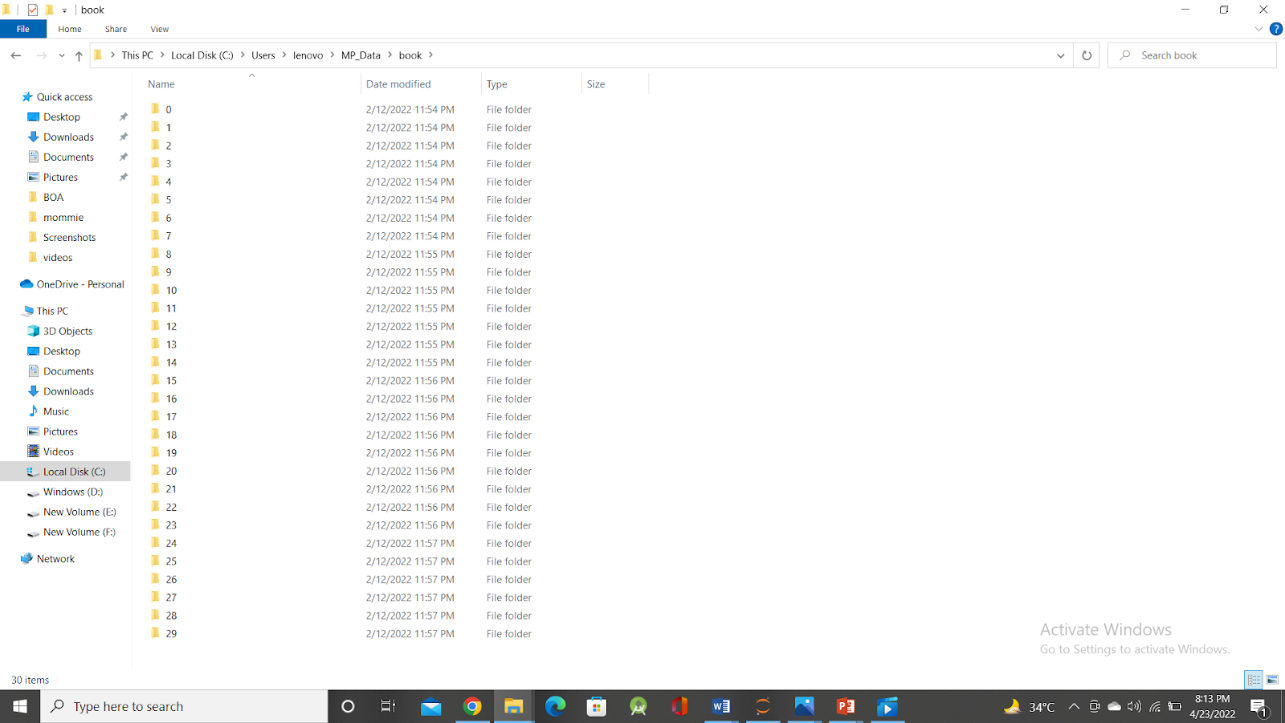


Fig – Dataset for Module 1

For the second module we stored different images of alphabets depicted in sign language and gifs of some commonly spoken words.

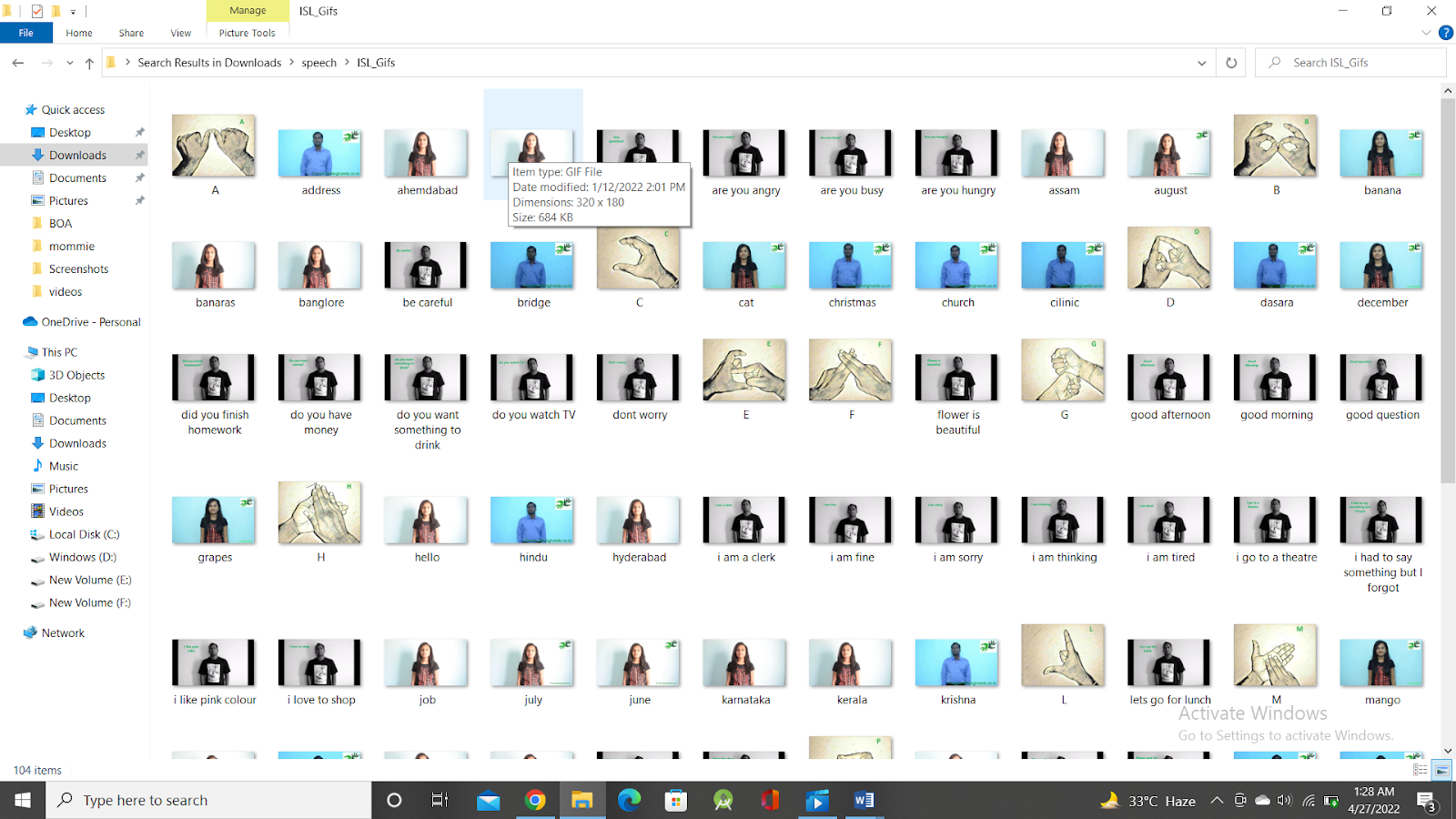


Fig – Dataset for Module 2 (i)

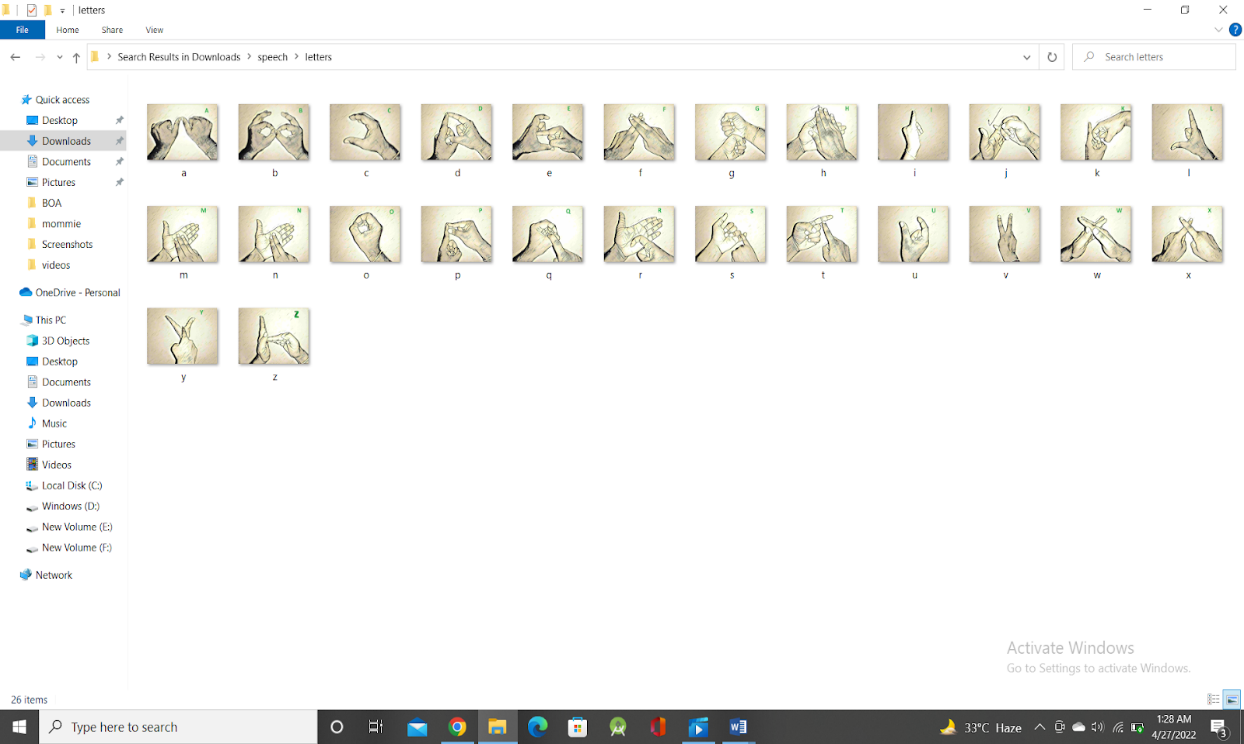


Fig – Dataset for Module 2 (ii)

**DESIGN METHODOLOGY AND ITS NOVELTY**

**4.1 Methodology and goal**

The entire project is divided into two modules namely Signify and Amplify.

1.     Signify- The overall goal of this module is to implement a real time sign language detection using **Long Short-Term Memory** (LSTM) neural network model built on tensor flow keras.

Siriak, R., Skarga-Bandurova, I., & Boltov, Y. (2019, September). Deep convolutional network with long short-term memory layers for dynamic gesture recognition. In *2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)* (Vol. 1, pp. 158-162). IEEE.

2.     Amplify- The overall aim of this module is to build an application which takes in live speech or audio recording as input, converts it into text and displays the relevant Indian Sign Language images or GIFs, using Natural Language Processing and Machine Learning Algorithm.

Dhanjal, A. S., & Singh, W. (2021). An automatic machine translation system for multi-lingual speech to Indian sign language. *Multimedia Tools and Applications*, 1-39.

**4.2  Functional modules design and analysis**

***4.2.1*  Module 1: Signify**

 To leverage a key point detection model to build a sequence of key points (i.e. using multiple frames for detection) which can then be passed to an action detection model to decode sign language.

 The key points will be extracted using the Media Pipe Holistic from the Hands, Face & Body. As part of the model building process, we will be using TensorFlow and Keras to build a deep neural network that leverages LSTM layers to handle the sequence of key points to make predictions in real time.

**Media Pipe Holistic:**

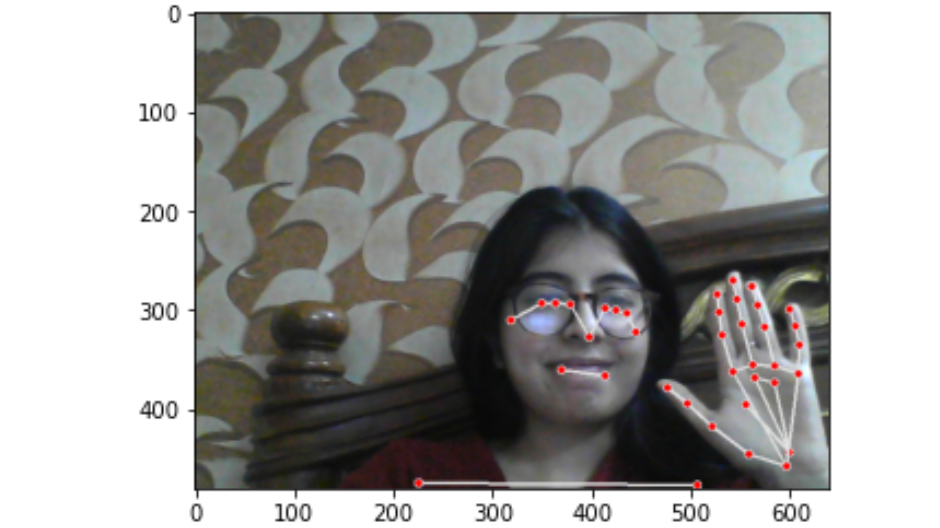


Fig- Key-point Detection via Media Pipe Holistic

Media-Pipe is an open-source, cross-platform Machine Learning framework used for building complex and multimodal applied machine learning pipelines. It can be used to make advanced machine learning models such as face detection, multi-hand tracking, object detection, and tracking, and much more.

Flutura, S. (2021). Mobile social signal interpretation in the wild for wellbeing.

Media-Pipe basically acts as a mediator for handling the implementation of models for systems operating in any field that assists an engineer focusing more on testing models, rather than the system.

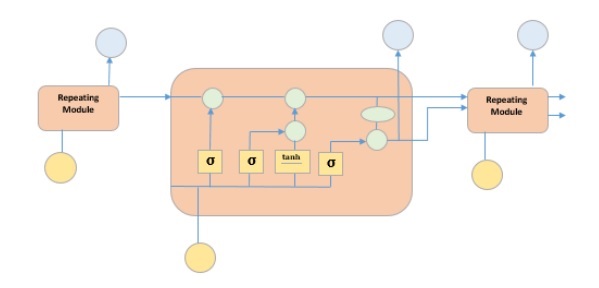
Media-pipe Holistic is one of the pipelines which contains optimised face, hands, and pose components which allows for holistic tracking, thus enabling the model to simultaneously detect hand and body poses along with face landmarks. One of the main usages of Media-Pipe holistic is to detect face and hands and extract key points to pass on to a computer vision model.

Maman, B., & Bermano, A. (2022). Typenet: Towards Camera Enabled Touch Typing on Flat Surfaces Through Self-Refinement. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 1140-1149).

**LSTM Model**

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network capable of learning order dependence in sequence prediction problems. This is a behaviour required in complex problem domains like machine translation, speech recognition, and more.

Veeriah, V., Zhuang, N., & Qi, G. J. (2015). Differential recurrent neural networks for action recognition. In *Proceedings of the IEEE international conference on computer vision* (pp. 4041-4049).



` Fig : LSTM Model

The picture above depicts four neural network layers in yellow boxes, point wise operators in green circles, input in yellow circles and cell state in blue circles. An LSTM module has a cell state and three gates which provides them with the power to selectively learn, unlearn or retain information from each of the units. The cell state in LSTM helps the information to flow through the units without being altered by allowing only a few linear interactions. Each unit has an input, output and a forget gate which can add or remove the information to the cell state. The forget gate decides which information from the previous cell state should be forgotten for which it uses a sigmoid function. The input gate controls the information flow to the current cell state using a pointwise multiplication operation of ‘sigmoid’ and ‘tanh’ respectively. Finally, the output gate decides which information should be passed on to the next hidden state.

***4.2.2 Module 2: Amplify***

An application which takes in live speech or audio recording as input, converts it into text and displays the relevant Indian Sign Language images or GIFs.

* Front-end is achieved using EasyGui.
* Speech is taken as input through a microphone using PyAudio.
* Speech recognition is done using Google Speech API and Sphinx(for offline use).
* Text Preprocessing was completed using NLP.
* Dictionary based Machine Translation.

*Algorithm-*

1. Get started
2. Finding a Speech
3. Listen for 1 second and measure the power limit of ambient sound levels.
4. Listen to the talk using the microphone. Now the power limit is already set to a good value, and we can hold the conversation honestly right away.
5. Know Speech.
6. Convert Speech to Text.
7. Make the text in lower case for continuous use.
8. Detected Text
   1. If “goodbye” then exit. 2. Else if Detected Text in predefined Dictionary Words. Display respective GIFs of the Phrase.
   2. Else Count the Letters of the Word/Phrase

i. Display the Visual of the phrase with some delay of Actions.

1. Continue all the steps from Step 3, and continue till the Speech Ends.
2. If Error in Step 2, that is if no Speech Detected then display error message “Could not listen”.

***EasyGUI***

EasyGUI is a very simple & easy GUI module for Python. EasyGUI differs from other GUI generators in that EasyGUI is NOT event-driven. Instead, all GUI interactions are required for simple operating calls. Unlike other complex GUIs, EasyGUI  is the simplest GUI to date.

Lambert, K. A. (2018). *Fundamentals of Python: first programs*. Cengage Learning.

Install using this command: pip install easygui

It is not recommended to use EasyGui in IDLE as EasyGui works in Tkinter and has its own event loop, and IDLE is also an application written in the Tkinter module and has its own event loop. So when both are run at the same time, conflicts can also have unintended consequences.

It is therefore preferable to use EasyGui without IDLE.

We used EasyGui to create a user interface which drives our model forward.

***Google Speech API***

Google has a great Speech Recognition API. This API converts spoken text (microphone) into written text (Python strings), briefly Speech to Text.

You can just simply speak in a microphone and Google API will translate this into written text. The API has excellent results for the English language.

A speech recognition API offloads the logic, such that you can simply send a web request to the API, which then returns the text that was recognized. You can do this from Python code directly, but your script will need internet access behind the scenes. This program will record audio from your microphone, send it to the speech API and return a Python string. The audio is recorded using the speech recognition module, the module will be included on top of the program. Secondly we send the record speech to the Google speech recognition API which will then return the output.

***PyAudio***

PyAudio provides Python bindings for PortAudio, the cross-platform audio I/O library. With PyAudio, we can easily use Python to play and record audio on a variety of platforms.

Kumari, D., Pai, N., & Nayak, P. (2017). Voice-based E-mail system for blinds. *NM AM INSTITUTE OF TE C HNOLOGY*.

PyAudio is basically inspired by:

•pyPortAudio/fastaudio: Python bindings for PortAudio v18 API.

•tkSnack: cross-platform sound toolkit for Tcl/Tk and Python.

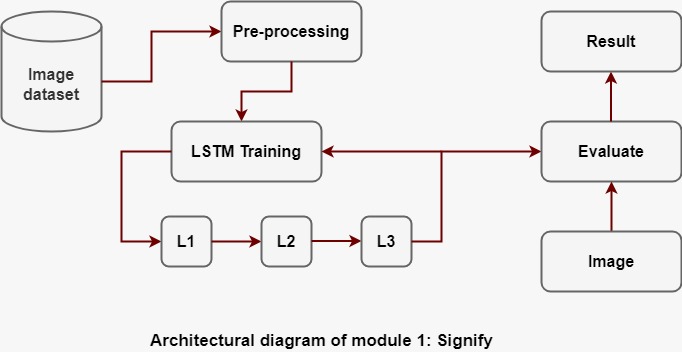
To use PyAudio, we first set up PyAudio using the pyaudio.PyAudio () function, which configures the portaudio system. To record or play audio, we can directly open the stream to the device we want with the desired audio parameters using pyaudio.PyAudio.open ().

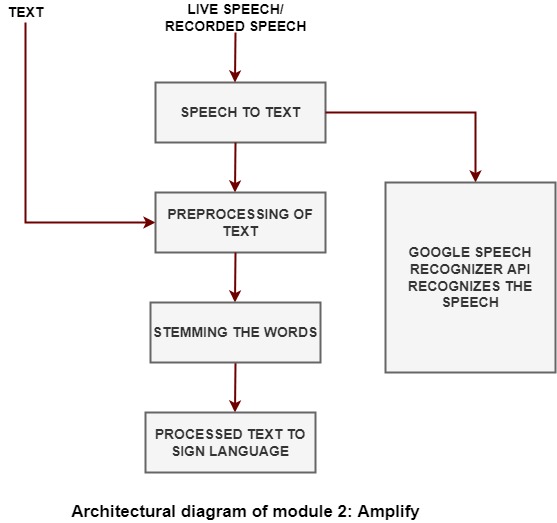
Wickert, M. (2018, July). Real-Time Digital Signal Processing using pyaudio\_helper and the ipywidgets. In *Proceedings of the 17th Python in Science Conference, Austin, TX, USA* (pp. 9-15).

This sets up pyaudio.Stream to play or record audio. Play audio by recording audio data from the stream using pyaudio.Stream.write (), and read audio data from the stream using pyaudio.Stream.read (). Use pyaudio.Stream.stop\_stream () to stop playing / recording, and pyaudio.Stream.close () to stop streaming.

Finally, terminate the portaudio session using pyaudio.PyAudio.terminate ().

***4.3 Software Architecture Diagram***





**4.4 User Interface designs**

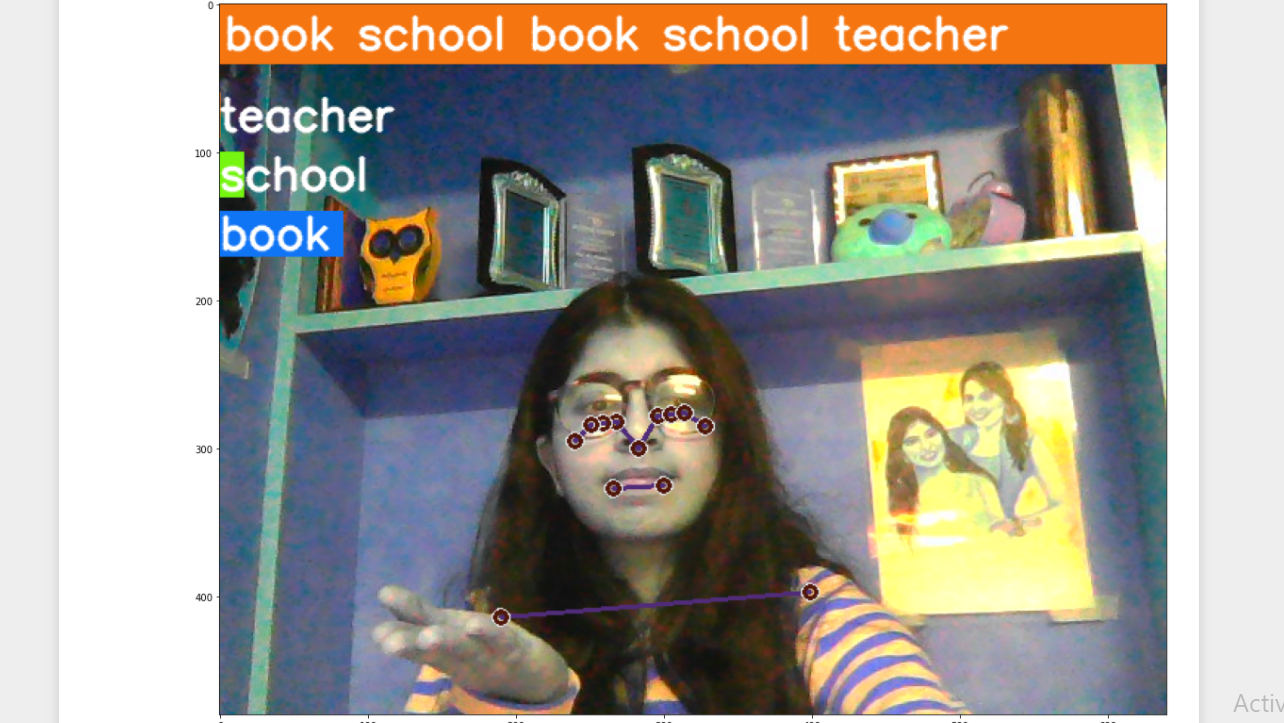
****

                               Fig – UI of Signify (Module 1)

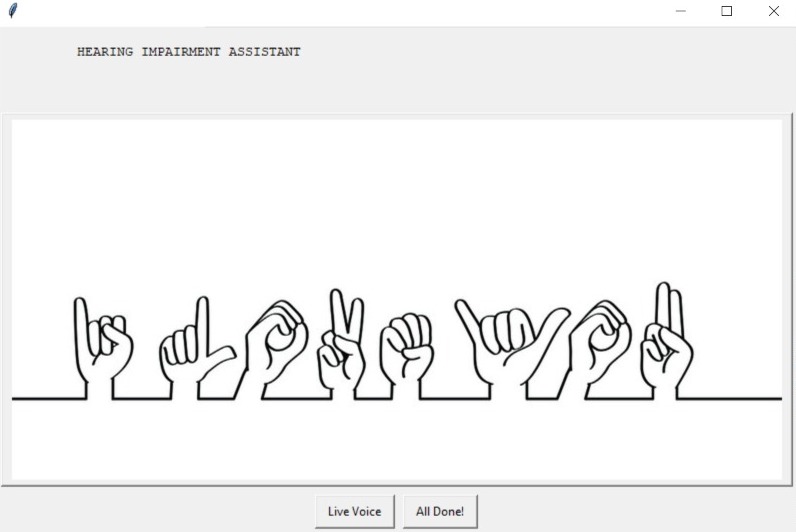


Fig- UI of Module 2 (Amplify)