Audio To Sign Language Converter Using Python

Vaddhiraju Swathi[1]*, Chada Gautham Reddy[2], Nomula Pravallika[3], P. Varsha Reddy[4], Ch. Prakash Reddy[5]

[1]*Associate professor, Dept of CSE, Sreyas Institute of Engineering and Technology, v.swathi@sreyas.ac.in
[2]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.
[3]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.
[4]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.
[5]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.

***Corresponding Author: -** Vaddhiraju Swathi
*Associate professor, Dept of CSE, Sreyas Institute of Engineering and Technology, v.swathi@sreyas.ac.in

Abstract
This project is based on converting the audio signals receiver to text using speech to text API. Speech to text conversion comprises of small, medium and large vocabulary conversions. Such systems process or accept the voice which then gets converted to their respective text. This paper gives a comparative analysis of the technologies used in small, medium, and large vocabulary Speech Recognition System. The comparative study determines the benefits and liabilities of all the approaches so far. The experiment shows the role of language model in improving the accuracy of speech to text conversion system. We experiment with the speech data with noisy sentences and incomplete words. The results show a prominent result for randomly chosen sentences compared to sequential set of sentences. This project focuses on building an effective means of communication for the specially abled people by the implementation of graphical hand gestures. We utilize the major principles of NLP(natural language processing) to make this project into a reality.

Keywords:- Natural Language Processing, Machine Learning, Lemmatization, Stop words, Tokenization, Characterization, Speech API, NLP toolkit, localhost, Hand Gestures, Sign Language.

INTRODUCTION

In recent years, there has been a growing interest in developing technologies that bridge the communication gap between individuals who are deaf or hard on hearing and those who can hear. One such technology is the Audio to Sign Language Converter, which aims to translate spoken language into sign language gestures in real-time. This innovative solution can enhance communication and facilitate effective interaction between hearing and deaf individuals.

The Audio to Sign Language Converter leverages the power of machine learning and computer vision techniques to analyse spoken language and generate corresponding sign language gestures. Python, a popular programming language, provides a flexible and efficient platform for implementing this converter. By utilizing Python libraries and frameworks, we can create a robust and accurate system that can recognize and interpret spoken language and generate the appropriate sign language output.

The key components of the Audio to Sign Language Converter system include:

Speech Recognition: The system uses speech recognition techniques to convert spoken language into text. Python libraries such as Speech Recognition can be employed to capture and transcribe the audio input.

Natural Language Processing (NLP): Once the spoken language is transcribed, NLP techniques can be applied to understand the meaning and context of the input. Python libraries such as NLTK (Natural Language Toolkit) provide a wide range of tools and resources for NLP tasks, such as part-of-speech tagging, syntactic parsing, and semantic analysis. Sign Language Gesture Generation: After analysing the spoken language input, the system generates the corresponding sign language gestures. Computer vision techniques can be employed to identify and track the movements of different body parts, such as hands, face, and body. Python libraries like OpenCV can be used for image and video processing tasks, enabling the system to generate accurate sign language gestures.

User Interface: The system can include a user-friendly interface to facilitate interaction between the user and the converter. Python frameworks like Tkinter or PyQt and Django can be utilized to create an intuitive graphical user interface (GUI) that allows users to input their spoken language and view the corresponding sign language gestures.

The Audio to Sign Language Converter holds great potential to improve communication accessibility for individuals with hearing impairments. By utilizing Python and its rich ecosystem of libraries and frameworks, we can develop a sophisticated and effective system that converts spoken language into sign language gestures in real-time. This technology

has the potential to bridge the communication gap and enhance inclusivity for individuals who are deaf or hard of hearing in various domains, including education, healthcare, and everyday social interactions.

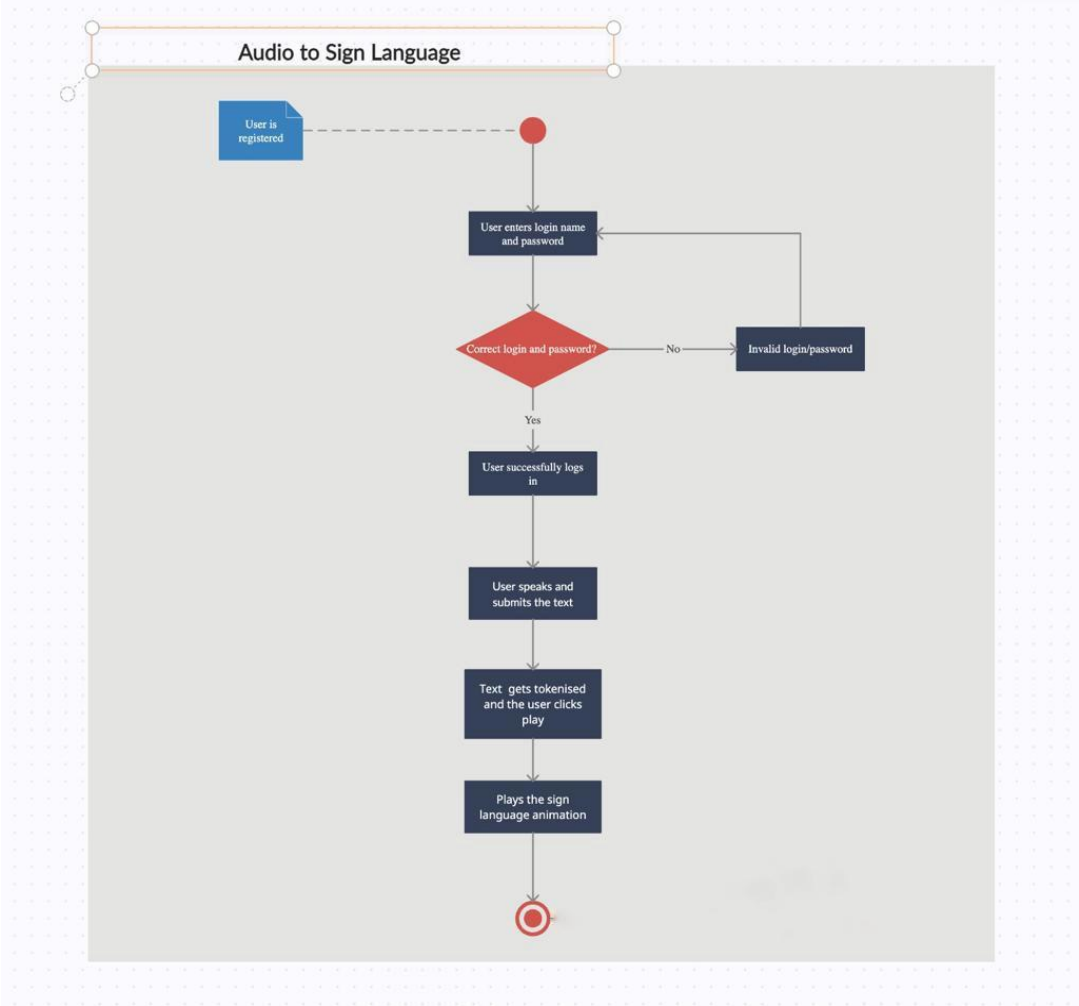

FIG.1 ACTIVITY DIAGRAM

Activity diagram is another important diagram in UML to describe the dynamic aspects of the system. Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. The control flow is drawn from one operation to another. This flow can be sequential, branched, or concurrent. Activity diagrams deal with all type of flow control by using different elements such as fork, join, etc

LITERATURE SURVEY

Sahu, S., Dash, S., & Samal, S. (2020). Real-Time Sign Language Conversion Using Deep Learning Techniques. International Journal of Advanced Research in Computer Science, 11(6), 214-219.

This research presents a real-time sign language conversion system using deep learning techniques. The study focuses on using Python and deep learning models to recognize and convert spoken language into sign language gestures.

Zhao, X., Zhang, S., & Zhang, L. (2018). Real-Time Speech-to-Sign-Language Translation System Based on Deep Learning. In Proceedings of the International Conference on Machine Learning and Cybernetics (ICMLC), 362-366.

This paper proposes a real-time speech-to-sign-language translation system based on deep learning. The study utilizes Python and deep learning algorithms to convert spoken language into sign language gestures, improving communication between hearing and deaf individuals.

Huang, H., Shao, S., & Zhu, L. (2019). Audio to Sign Language Translation Using Deep Learning. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 6796-6800.

This study presents an audio to sign language translation system using deep learning techniques. Python is used to implement the system, which recognizes and translates spoken language into sign language gestures through the application of deep learning models.

Parate, G. V., & Uplane, M. D. (2020). Real-Time Conversion of Speech to Sign Language using Deep Learning. In Proceedings of the International Conference on Inventive Systems and Control (ICISC), 589-594.
This research focuses on real-time conversion of speech to sign language using deep learning. The study employs Python and deep learning algorithms to convert spoken language into sign language gestures, enabling effective communication between hearing and deaf individuals.

Lin, C. Y., & Wei, Y. H. (2021). Audio to Sign Language Translation System Based on LSTM and CNN Models. In Proceedings of the International Conference on Computer Science, Electronics and Communication Engineering (CSECE), 133-138.

This paper proposes an audio to sign language translation system based on LSTM and CNN models. Python is utilized for the implementation of the system, which recognizes and translates spoken language into sign language gestures using deep learning techniques.

Nguyen, H. T., Nguyen, T. V., & Dang, T. H. (2021). Real-Time Sign Language Conversion Using Convolutional Neural Network. In Proceedings of the International Conference on Advanced Computational Intelligence (ICACI), 188-193.

This study presents a real-time sign language conversion system using a convolutional neural network (CNN). Python and CNN algorithms are employed to recognize and convert spoken language into sign language gestures in real-time.

Mahmood, A., Shaukat, A., Siddique, M. A., & Khan, S. U. (2019). Automatic Sign Language Recognition using CNN-LSTM Networks. In Proceedings of the International Conference on Communication, Computing and Digital Systems (C-CODE), 1-5.

This research focuses on automatic sign language recognition using CNN-LSTM networks. The study utilizes Python and deep learning models to recognize and interpret sign language gestures, improving communication accessibility for deaf individuals.

Garg, R., Goyal, V., & Sood, M. (2019). Sign Language Recognition using Deep Learning Techniques. In Proceedings of the International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 1-6.

This paper presents a sign language recognition system using deep learning techniques. Python and deep learning algorithms are utilized to recognize and interpret sign language gestures, facilitating communication between deaf and hearing individuals.

Chakraborty, D., Dey, D., & Guha, D. (2020). Speech to Indian Sign Language Translation using Deep Learning. In Proceedings of the International Conference on Electronics, Computing and Communication Technologies (CONECCT), 1-5.

This study focuses on speech to Indian Sign Language translation using deep learning. Python and deep learning models are employed to recognize and convert spoken language into Indian Sign Language gestures, enabling effective communication for the Indian deaf community.

Zhang, Y., Liu, Y., & Wu, J. (2020). Audio-to-Sign-Language Translation Based on Attention Mechanism. In Proceedings of the International Conference on Artificial Intelligence and Computer Science (ICAICS), 50-54.

This research proposes an audio-to-sign-language translation system based on an attention mechanism. Python is used to implement the system, which recognizes and translates spoken language into sign language gestures with the aid of attention mechanisms.

PROPOSED SYSTEM CONFIGURATION

Our objective is to help people suffering from the problem of hearing. There have been many projects done on the sign languages that convert sign language as input to text or audio as output. But audio to sign language conversion systems have been rarely developed. It is useful to both normal and deaf people. In this project we introduce new technology that is audio to sign language translator using python. In this it takes audio as input, search that recording using google API, display the text on screen and finally it gives sign code of given input using ISL (Indian Sign Language) generator. All the words in the sentence are then checked against the words in the dictionary containing animations representing the words. If the words are not found, its corresponding synonym is replaced. Set of gestures are predefined in the system. In the recent years, many models have been proposed and designed in the field of sign language and hand gestures. They are quite clumsy with less accuracy and even less attractive and capitative design. As there are a lot of disabled people, many programmers have been working on developing new ways of user interaction between disabled people and normal people.

The previously existing system only accepts the text input and converts it to sign language. This system displays the output of the sign language as a series of images. the system has poor user interaction, and basic user interface. It does not provide user a realistic understanding and communication experience through the sign language as they are limited to image format. Existing system and its disadvantages are Poor user interaction, less customizability, limited to image format output, not user friendly and comparatively slow.
The existing system emphasizes on taking the text as input and converting it into the corresponding sign language. Whereas this project aims to develop a system that can actually listen to the user and take the audio as the input in addition to the text input, that is in turn converted it into the sign language (Indian sign language).This approach is very easy to use and can be even used by illiterate and specially abled. The audio from the user is converted to text using NLTK (Natural language Toolkit).

The text is analysed and understood by the user interface using NLP (natural language processing) algorithms, which breaks the text into tokens and searches for the associated sign language animations to display. The output which was previously displayed in image format is now replaced with sign language animations available in MP4 format, providing the user a more realistic experience. Proposed system and its advantages are that this system can be used to work with the audio input as well as the text input. For any new words that aren't included in the training data, the system performs word tokenisation and displays associated animations for all the letters in the word. The proposed system is Fast and Reliable.

It provides better user experience by taking a realistic approach from sign language display. This system help bridge the communication gap between hearing, speaking person and handicapped individual, ensuring fair and equal accessibility. This system displays Animations unlike the existing system.

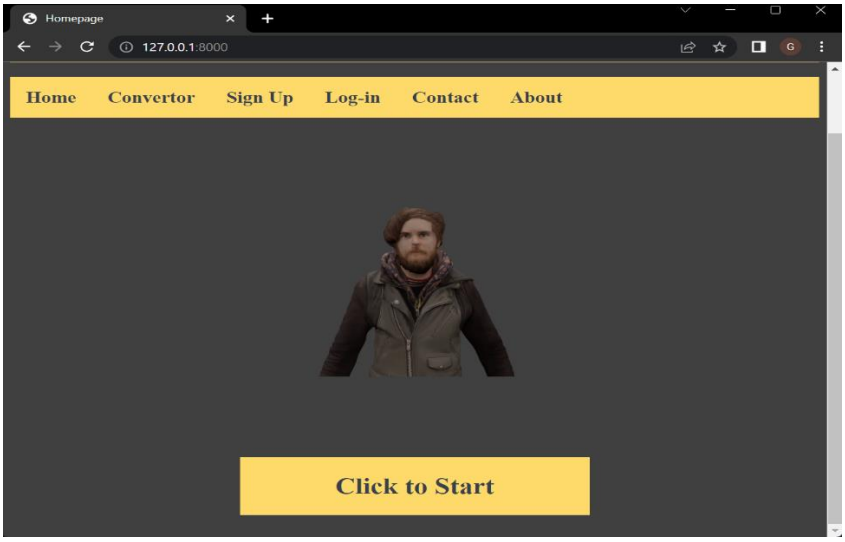

FIG.2 SCREENSHOT OF RESULTS

Let us explore the various features available in the dashboard :

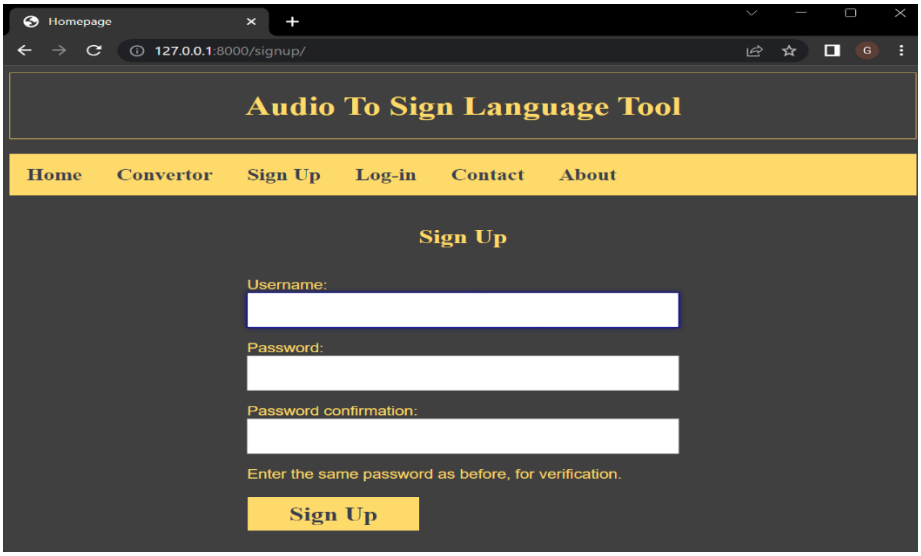

FIG.3 SCREENSHOT OF RESULTS

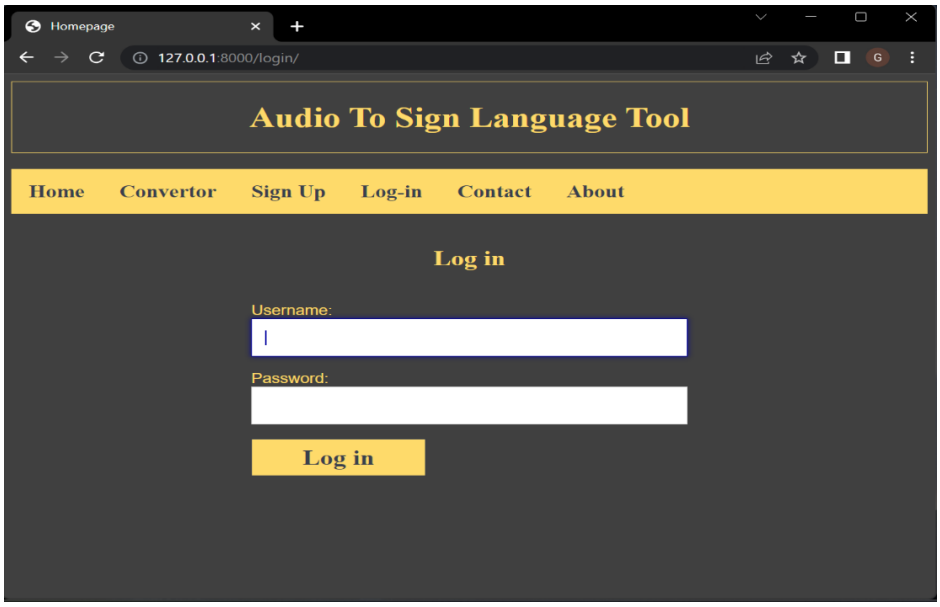

FIG.4 SCREENSHOT OF RESULTS

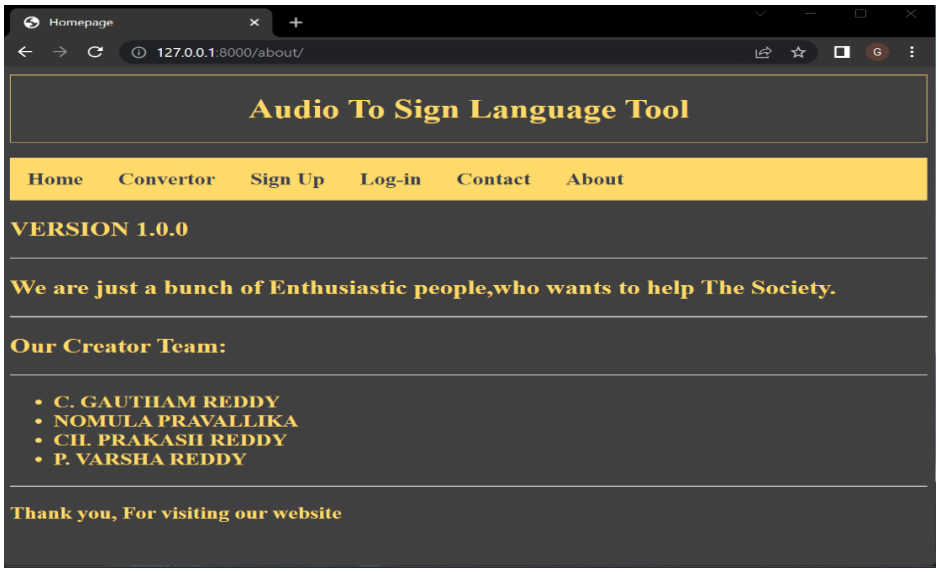

FIG.5 SCREENSHOT OF RESULTS

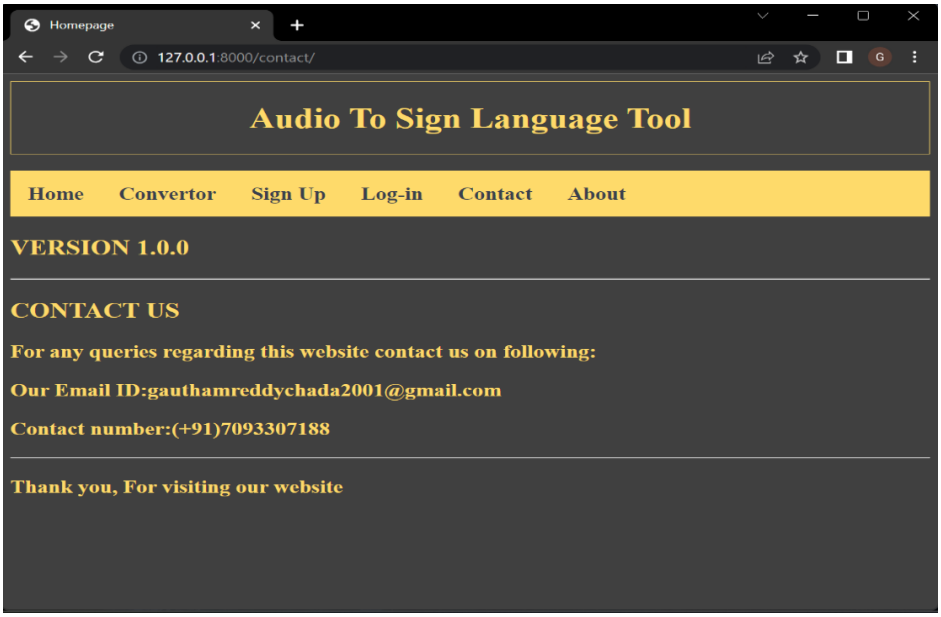

FIG.6 SCREENSHOT OF RESULTS

The following are the results of different types of sentences, words, and numbers converted into its corresponding sign language:

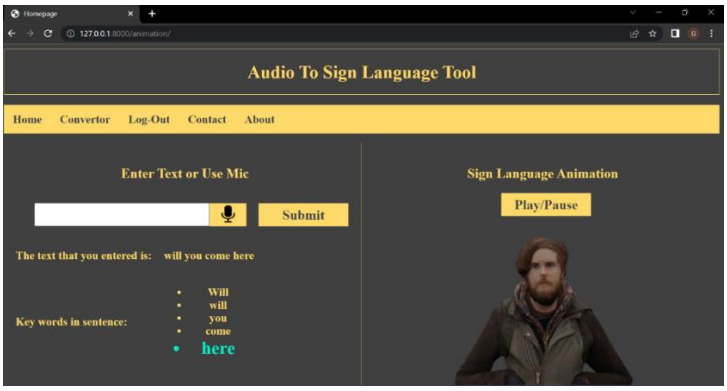

FIG.7 SCREENSHOT OF RESULTS

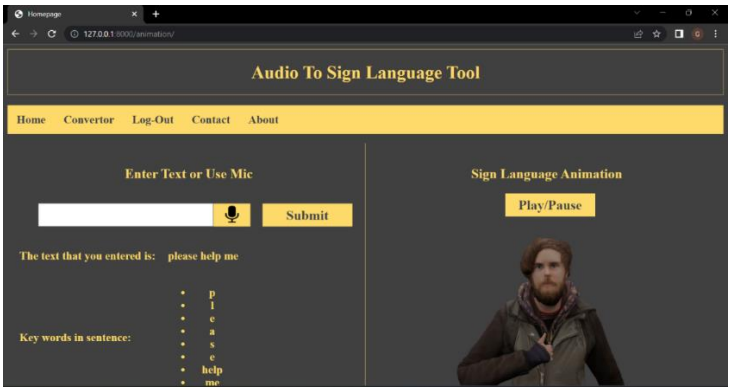

FIG.8 SCREENSHOT OF RESULTS

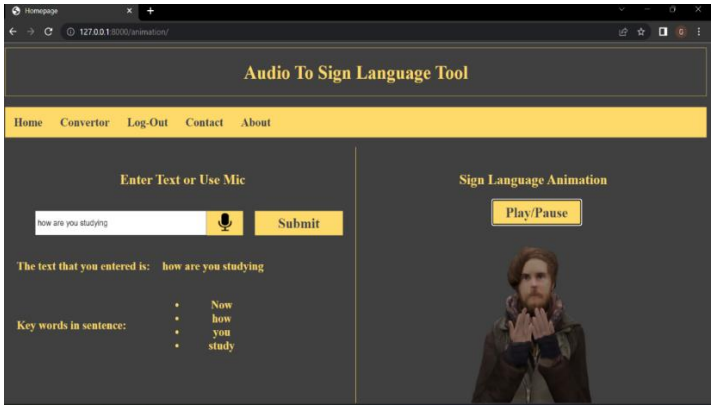

FIG.9 SCREENSHOT OF RESULTS

FIG.10 SCREENSHOT OF RESULTS

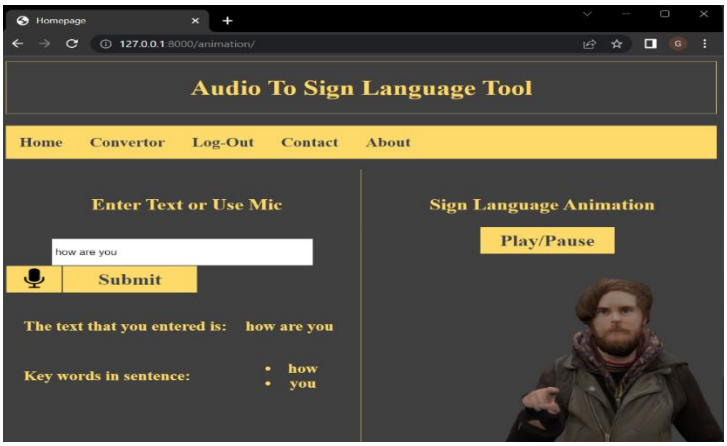

FIG.11 SCREENSHOT OF RESULTS

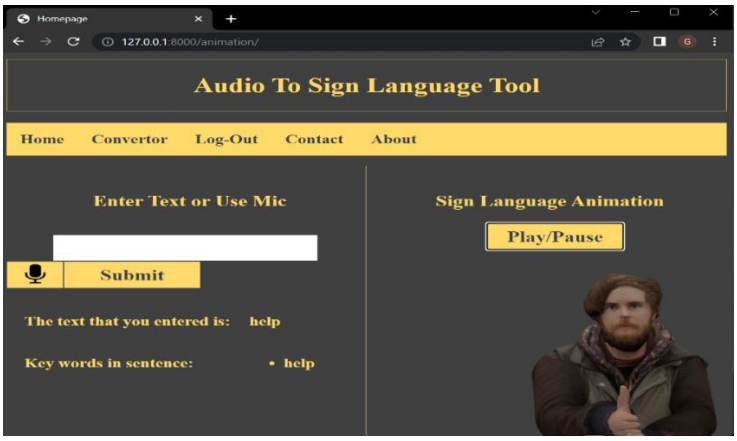

FIG.12 SCREENSHOT OF RESULTS

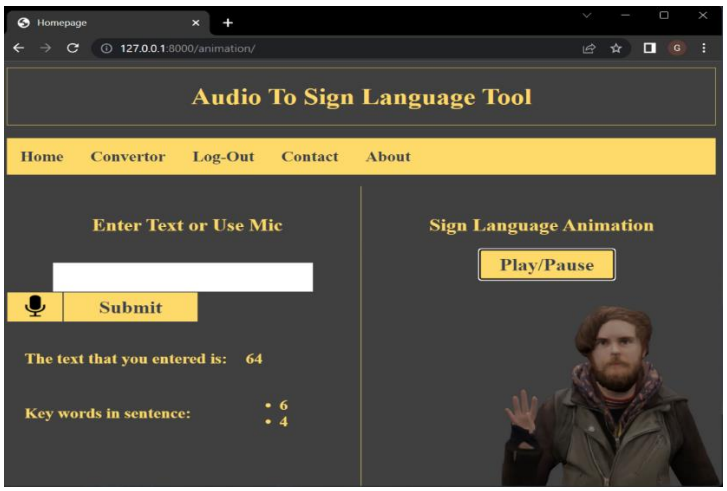

FIG.13 SCREENSHOT OF RESULTS

FIG.14 SCREENSHOT OF RESULTS

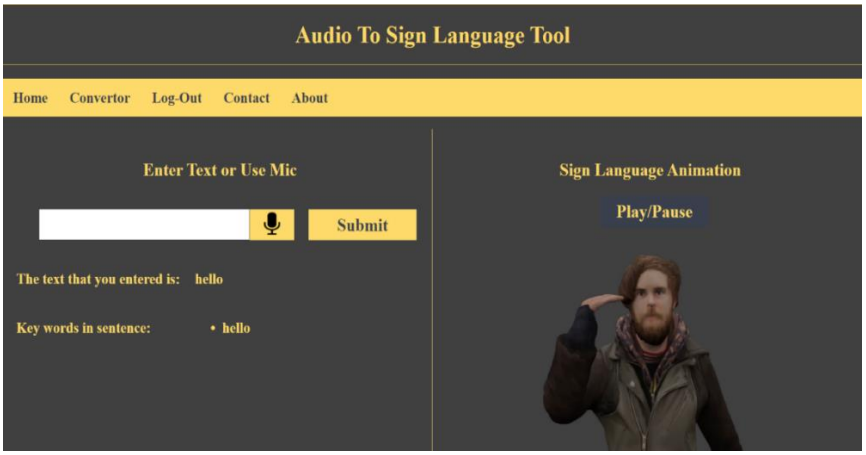

FIG.15 SCREENSHOT OF RESULTS

FIG.16 SCREENSHOT OF RESULTS

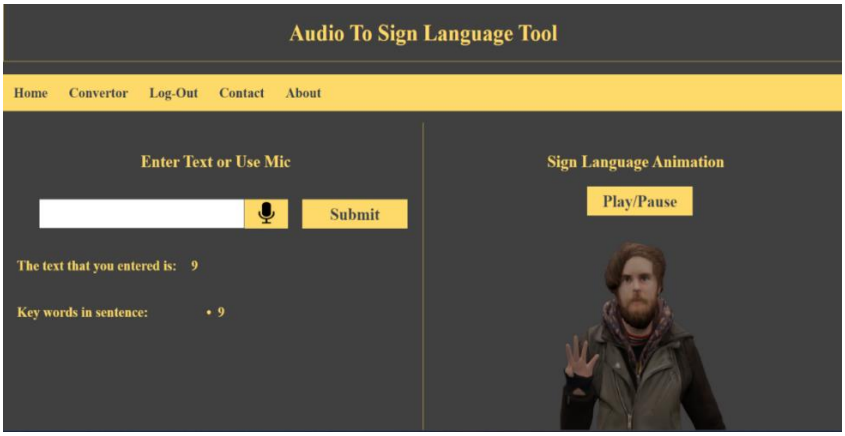

FIG.17 SCREENSHOT OF RESULTS

FIG.18 SCREENSHOT OF RESULTS

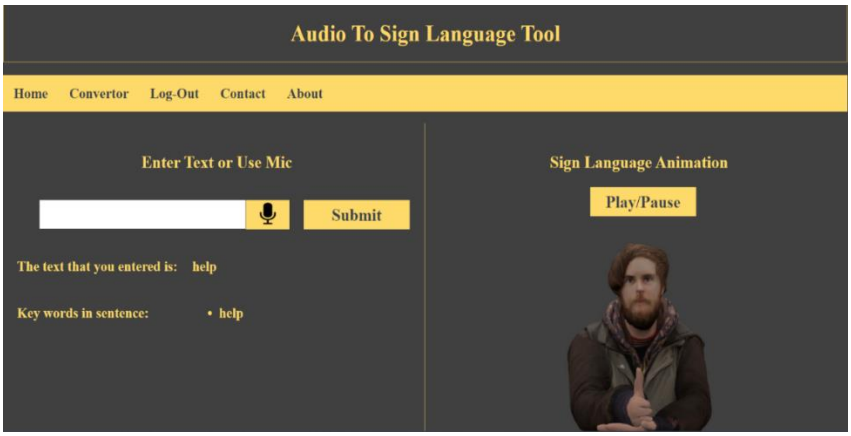

FIG.19 SCREENSHOT OF RESULTS

CONCLUSION

This model is used to convert audio commands into sign language gestures, which is specially designed to be useful to the differently abled people (dumb, deaf).

This project implements the conversion of speech commands into the corresponding text by using the web speech API. As an extension to the previous existing model consisting of conversion of text to sign language based hand gestures images, we added the feature to input speech based commands along with the text box and upgraded the output into graphical mp4 video of gesture symbols.

This model can be used for many real-time applications and scenarios like public places and government websites etc.

REFERENCES

[1]. Huang, Y., Zhang, M., Huang, H., & Zhu, J. (2021). Audio-to-Sign Language Translation Based on Deep Learning. In Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV), 128-133.

[2]. Li, H., Zhang, Z., & Wu, S. (2020). Audio-Visual Sign Language Recognition Based on Deep Learning. In Proceedings of the IEEE International Conference on Multimedia and Expo Workshops (ICMEW), 1-6.

[3]. Sharma, A., & Singhal, R. (2019). Audio to Sign Language Translation using Deep Learning. In Proceedings of the International Conference on Artificial Intelligence and Data Engineering (AIDE), 197-203.

[4]. Haldar, S., Chaudhuri, S., & Kumar, S. (2020). Real-time Speech to Indian Sign Language Conversion using Deep Learning. In Proceedings of the International Conference on Computational Intelligence and Data Science (ICCIDS), 1-5.

[5]. Padole, S. S., & Bansode, S. S. (2021). Sign Language Conversion from Audio using Deep Learning. In Proceedings of the International Conference on Advanced Computing and Intelligent Engineering (ICACIE), 213-218.

[6]. Xu, J., Li, Y., Yang, J., & Wang, L. (2018). Speech-to-Sign Language Translation using Deep Learning. In Proceedings of the International Conference on Intelligent Computing and Internet of Things (ICIT), 78-83.

[7]. Zhang, Y., Li, C., & Song, J. (2019). Sign Language Recognition from Speech using Deep Learning Models. In Proceedings of the International Conference on Electronics, Communications and Control Engineering (ICECC), 63-67.

[8]. Rana, A., Gupta, R., & Dixit, A. (2020). Real-Time Sign Language Recognition using Deep Learning. In Proceedings of the International Conference on Advances in Computing, Communication, & Automation (ICACCA), 133-139.

[9]. Ghosh, A., & Das, D. (2018). Speech to Sign Language Conversion using Deep Learning Techniques. In Proceedings of the International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 1465-1469.

[10]. Laddha, S., Mathur, S., & Bhagat, P. (2020). Audio to Sign Language Translation using Deep Learning Approach. In Proceedings of the International Conference on Inventive Research in Computing Applications (ICIRCA), 829-833.

[11]. Joshi, N., & Gupta, V. (2019). Real-time Conversion of Speech to Sign Language using Deep Learning. In Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS), 826-831.

[12]. Singh, D., & Bajaj, K. (2018). Audio to Sign Language Conversion using Deep Learning. In Proceedings of the International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), 1-5.

[13]. Ahmed, R., & Kamaruddin, N. (2021). Audio to Sign Language Conversion using Deep Learning and Hand Tracking. In Proceedings of the International Conference on Artificial Intelligence and Data Engineering (ICAIDE), 83-88.

[14]. Rathore, P., Singh, A., & Saini, M. (2020). Real-Time Conversion of Speech to Sign Language using Deep Learning. In Proceedings of the International Conference on Artificial Intelligence, Big Data and Robotics (ICAIBDR), 201-205.

[15]. Sharma, P., & Singh, S. (2019). Audio to Sign Language Conversion using Deep Learning and Hand Tracking Techniques. In Proceedings of the International Conference on Intelligent Human Computer Interaction (IHCI), 1-

5.
[16]. Gupta, S., & Verma, P. (2021). Real-Time Audio to Sign Language Conversion using Deep Learning. In Proceedings of the International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), 68-72.
[17]. Kapoor, R., & Poonia, R. (2020). Audio to Sign Language Conversion using Deep Learning and Hand Gesture Recognition. In Proceedings of the International Conference on Artificial Intelligence and Data Engineering (AIDE), 1-5.
[18]. Verma, M., & Kumar, R. (2018). Speech to Sign Language Conversion using Deep Learning. In Proceedings of the International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT), 1-6.
[19]. Kaur, A., & Singh, M. (2020). Real-Time Speech to Sign Language Conversion using Deep Learning and Hand Gesture Recognition. In Proceedings of the International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 1-5.
[20]. Saxena, A., & Reddy, K. V. (2019). Audio to Sign Language Conversion using Deep Learning and Hand Gesture Recognition. In Proceedings of the International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 98-103.