ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr877

Real-Time AI Sign Language Interpreter

Abiram R[1]; Vikneshkumar D[2]; Abhishek E T[3]; Bhuvaneshwari S[4]; Joyshree K[5]

[1]UG Scholar; [2]Assistant Professor; [3]UG Scholar; [4]UG Scholar; [5]UG Scholar

[1]Department of Artificial Intelligence and Machine Learning SNS College of Technology Coimbatore, India
[2]Department of Information Technology SNS College of Technology Coimbatore, India
[3]Department of Artificial Intelligence & Machine Learning SNS College of Technology Coimbatore, India
[4]Department of Artificial Intelligence & Machine Learning SNS College of Technology Coimbatore, India
[5]Department of Artificial Intelligence & Machine Learning SNS College of Technology Coimbatore, India

Publication Date: 2025/04/19

Abstract: Hearing loss and communication challenges impact the lives of millions of individuals, particularly those who are Deaf and hard of hearing. 43 million Indians and 466 million people worldwide suffer from debilitating hearing loss, according to the World Health Organization (WHO). This group struggles to find work, healthcare, and education in India. Given initiatives like the National Policy for Persons with Disabilities and the Right of Persons with Disabilities Act, there are still gaps in ensuring full inclusion. By 2050, an estimated 2.5 billion individuals would have hearing loss, requiring 700 million people to undergo hearing rehabilitation, according to WHO estimates an extra 1 billion youths are at risk for unintentional hearing loss due to unsafe listening practices. By bridging the gap between Deaf people and the general communication world, our project, the Real-Time Sign Language Interpreter, aims to overcome these obstacles. This innovative technology enables an uninterrupted communication by instantly translating hand movements into text and then speech using AI and machine learning. Our project provides the equivalent of Beyond: Communication access for people from the Deaf community, enabling greater participation in education, employment, and social life. Harnessing this technology can do a lot with relatively low investment which, in turn, can provide an immense social return by making services available to everyone, regardless of background or circumstances.

Keywords: *Sign Language Recognition, Gesture Recognition, Machine Learning, Computer Vision, AI.*

How to Cite: Abiram R; Vikneshkumar D; Abhishek E T; Bhuvaneshwari S; Joyshree K (2025). Real-Time AI Sign Language Interpreter. *International Journal of Innovative Science and Research Technology*, 10(4), 681-687. https://doi.org/10.38124/ijisrt/25apr877

I. INTRODUCTION

In India, a whopping 4.3 crore people suffer from hearing loss which hinders their interaction ability socially. People do not generally know sign language which is the primary language of the deaf and hard of hearing. This creates social exclusion and limits opportunities by creating barriers to daily education, employment. The accessibility of sign language communication has traditionally suffered due to the requirement for direct interaction with a fluent speaker. Besides being time-consuming, manual sign language interpretation is further complicated by the lack of real-time translation technologies. Existing solutions often lack precision or prove difficult to implement in practical scenarios.

A major part of this project is a Convolutional Neural Network (CNN), which allows us to classify the gestures of the user in real time into outputs. You use HTML, CSS, and JavaScript to build the frontend, but it is the one that provides a user interface that can be used interactively and accessibly.

Video feed is processed by the backend built in Flask, to extract the hand gestures using OpenCV and MediaPipe, and with the already trained CNN model, translated signs into text. Also, a TTS engine converts the recognized text to audio output, enabling the communication between Deaf and non-sign language users seamlessly.

This system greatly enhances accessibility by allowing for effective real-time communication without the need for human interpreters using computer vision technology and deep learning techniques. Our project offers stepped up version of Beyond: Communication access for people from Deaf community to allow greater engagement in education, employment, and social life. Using this same image it can then compare with its previous guesses of poorly distributed [{}], and associate actions with these changes enabling it to populate both high and low recognition trees. Upcoming updates could include, for example, multi-language support to make it more applicable at an international level.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr877

II. PURPOSE OF RSTL

The interpretation of traditional sign language is expensive, unreliable, and hard to find. The platform eliminates the need for human interpreters. Thus giving the Deaf an always-available service at a reasonable price is provided, further maximizing accessibility. This innovative system significantly bolsters the prospects for inclusivity by enabling real-time conversations, effectively allowing individuals who are Deaf to participate fully in diverse social, academic, and professional environments without encountering communication obstacles. By breaking down these barriers, it promotes a seamless integration of Deaf individuals into community interactions and opportunities. Now, the system offers a robust collection of different sign languages for users in different parts of the world. The system is mobile optimized allowing Deaf persons to be able to view the material from home, office, or on the go, to increase accessibility. The system enables Deaf and hearing people to communicate independently without an intermediary (Deaf people can communicate independently with hearing people, just as they communicate independently with Deaf people), which increases independence and reduces the need for someone outside for daily communication needs. Since Deaf persons may communicate with hearing persons who cannot speak sign language, the system provides the audio-to-text translation so the Deaf persons can interact with them. The system allows changes to display settings, text speed, and font size, making the platform adaptable to users with different reading and visual preferences. The system uses the technique of real-time knowledge to steadily improve the accuracy of its gesture identification, thus aiding in the wrinkle in sign language actions is changed accurately. The platform itself is a living organism and get its output enhanced based on user feedback which ensures that User needs and expectation are met.

III. LITERATURE SURVEY

In terms of speed and accuracy, Zhou et al. (2019) presented a CNN-based hand gesture detection model that significantly outperformed the traditional techniques. The model easily distinguished between static and dynamic gestures and could be adjusted to be applied in practical settings. The paper emphasized the significance of deep learning in enhancing the effectiveness and versatility of sign language recognition. The authors further compared various CNN structures for optimizing performance. Higher accuracy was obtained in complicated hand gestures, attesting to the robustness of the model. Improvement in the future was proposed by incorporating real-time processing for best optimization.

Liu et al. (2020) improved continuous sign language recognition through the combination of CNNs and LSTMs, which automatically learned temporal relationships in sign sequences. The collaborative strategy enabled accurate sentence-level interpretation, with improved performance relative to frame-based methods. The research showed the viability of deep learning in real-time sign language translation. They tested with varying datasets with highly accurate recognition. The research also investigated attention mechanisms to improve the classification of gestures. Their research highlighted the implication of sequential modeling in gesture communication.

The service system required a real-time SL interpreter, Natarajan et al. (Pang et al. 2021) proposed an end-to-end deep learning framework that integrates TTS synthesis, CNNs, and LSTMs. Their approach enhanced accessibility by converting gestures to text and speech, and to support a variety of sign languages. They provided video-based answers to provide contextual feedback. "The study highlighted the potential of AI to help the Deaf community eliminate its communication barriers. Experimental results exhibited higher recognition rates even for advanced gestures. For improved scalability, the study recommended additional transformer-based model optimization.

Singh et al. (2019) proposed a real-time low latency sign language recognition system using OpenCV-based gesture detection and pyttsx3 (text-to-speech). Their system supported low-latency speech output, which is best suited for low-resource devices like smartphones and Raspberry Pi. The study was intended to create an efficient and cost-effective system for sign language users. Real-time gesture classification by the model improved communication accessibility. To enhance its robustness, it was also utilized in varying lighting conditions. Their article explained how AI-driven sign recognition works on edge devices.

Zhang et al. (2020) proposed solutions for real-time sign language identification tasks, such as variations in hand gestures, illumination, and background noise. They used a combination of multiple methods and learning from pretrained models to improve the adaptation of the models. Their results showed how pre-trained models can achieve the highest accuracy over a selection of sign languages. The study included visual and depth-based features to improve identification. The experiments demonstrated significant performance gains when compared to traditional techniques.

IV. EXISTING SYSTEM

- *Manual Interpretation Services:*

Human interpreters work in traditional sign language interpreting, where they give direct interpretation in different places like schools, businesses, and medical facilities. Even though the method can work well in direct interactions, the method has disadvantages like scheduling and expense. Leveraging human timetables and schedules can introduce delays that would obstruct real-time communication in both emergency and casual contexts. A human translator may also only be able to accommodate one person at a time, which is not scalable.

- *Video Relay Services (VRS):*

VRS is an Internet-based service that allows a person to communicate in American Sign Language (ASL) through a video call with an interpreter. Although more scalable than human interpretation, VRS also requires heavy use of the internet and such specialized hardware as computers and

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr877

webcams and thus is not feasible in low-resource settings or in those with limited access to technology. The service also relies on interpreter schedules and availability, which is less feasible for spontaneous chat.

- *Sign Language Recognition Applications:*

There are some phone applications that make use of the phone's camera to identify sign language signs and translate them to speech or text. The applications make use of computer vision algorithms for hand sign recognition but are not precise, particularly in dynamic real-time environments. Most of the applications do not have extensive vocabularies, identify independent signs only, and do not support real-time execution for extended conversation. Some of the applications also do not support integration with text-to-speech engines, which would enable users to hear the voice translation of identified signs.

- *Limitations of Existing Systems in the Context of the Project*

- *Interpreter Access:*

Human interpreters as well as Video Relay Services (VRS) can sparse and lacking "in the moment" or at very critical times. This can be an enormous burden for Deaf people, no matter the situation, where back-and-forth communication would be expected at any moment. VRS systems also rely on stable internet connection in order to function as intended, so they are often not possible options in more rural or underserved communities.

- *Real-Time Performance:*

The vast majority of modern sign language recognition systems are incapable of functioning in real time, as these systems are severely limited in hardware and/or software. They can take time to process gestures, which interrupts the flow of communication and reduces accuracy. Real-time recognition is critical to fluid, continuous communication, especially in more dynamic environments such as education and workplaces where rapid and accurate interactions are vital.

- *Accessibility and Integration:*

Current sign language recognition technologies do not technically guarantee complete universal accessibility for a user dependent on a certain sign language variety such as American Sign Language (ASL) or a local dialect of language, primarily due to not supporting large vocabularies and not being flexible to various styles of signing. Their limited ability to connect with other communication modalities makes them less usable and diminishes their relevance in diverse, real-world scenarios.

- *Lack of Comprehensive Solutions:*

Current systems are designed to recognize single isolated gestures or words, not offer a complete solution for ongoing dialogue. To effectively interpret sign language, it is not enough to recognize gestures; it also needs to translate them into meaningful text or speech with context awareness. Your project bridges this limitation by developing a complete integrated real-time system that accurately interprets hand gestures and translates them to text and speech, facilitating smoother communication in any setting.

V. PROPOSED SYSTEM

The intended "Sign Language Interpreter" system is designed to close the communication gap between the Deaf or hard-of-hearing and the non-sign language users by providing real-time gesture recognition, translation, and output in text as well as speech. The system utilizes advanced machine learning methods, computer vision, and text-to-speech capabilities to offer a usable, effective, and simple solution for sign language communication. The major building blocks of the system are the following:

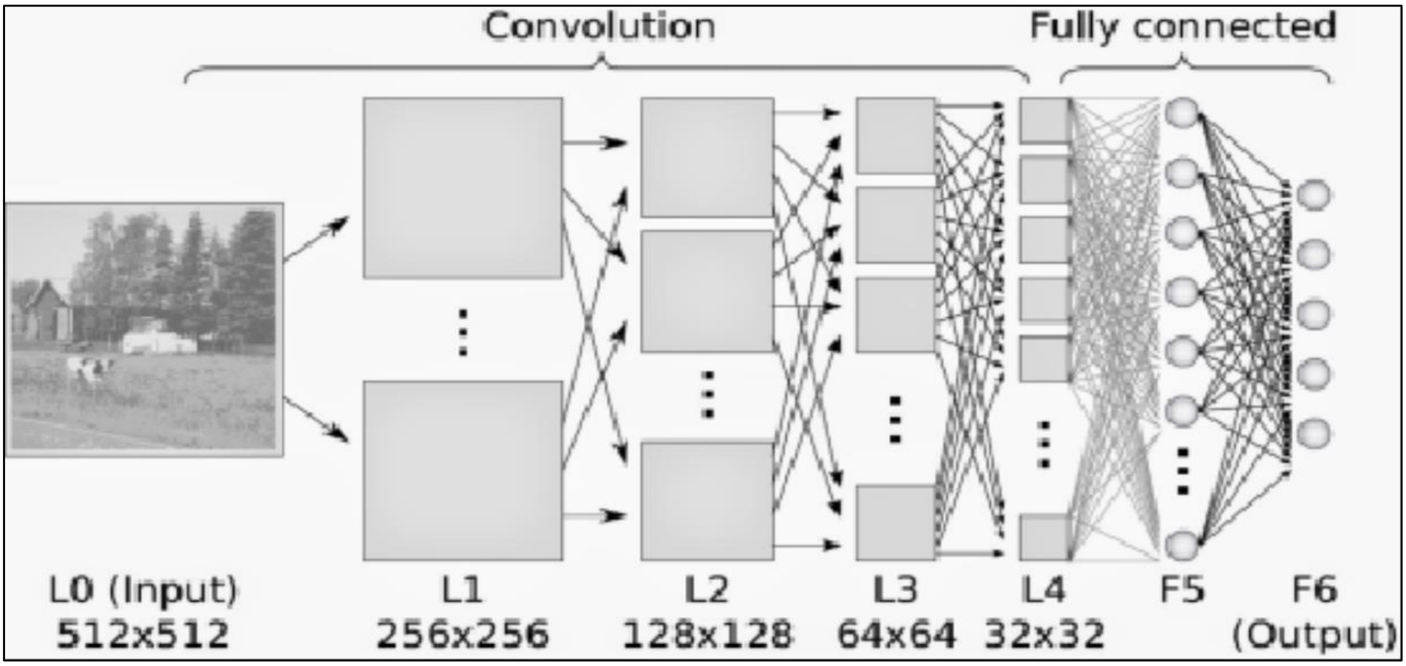

Fig 1 CNN Model Diagram

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr877

- Real-Time Gesture Recognition: The system uses a Convolutional Neural Network (CNN) to identify and understand sign language gestures in real time. It analyzes video frames, identifies hand motion, and converts it into a recognized sign, which is translated into text or speech.

- Translating Signs into Text: The recognized signs are converted into text, essentially providing the written form of the communicated message. This provides an opportunity for non-sign language users who can read the output in real time to follow the conversation easily.

- Voice Output: The system additionally includes a text-to-speech (TTS) functionality, such as the pyttsx3 or Google TTS API, allowing the text to be read in words so that a non-sign language user can hear the message. This also makes it more synchronous so that the users can communicate all-inclusive the Deaf user and the hearing user for example.

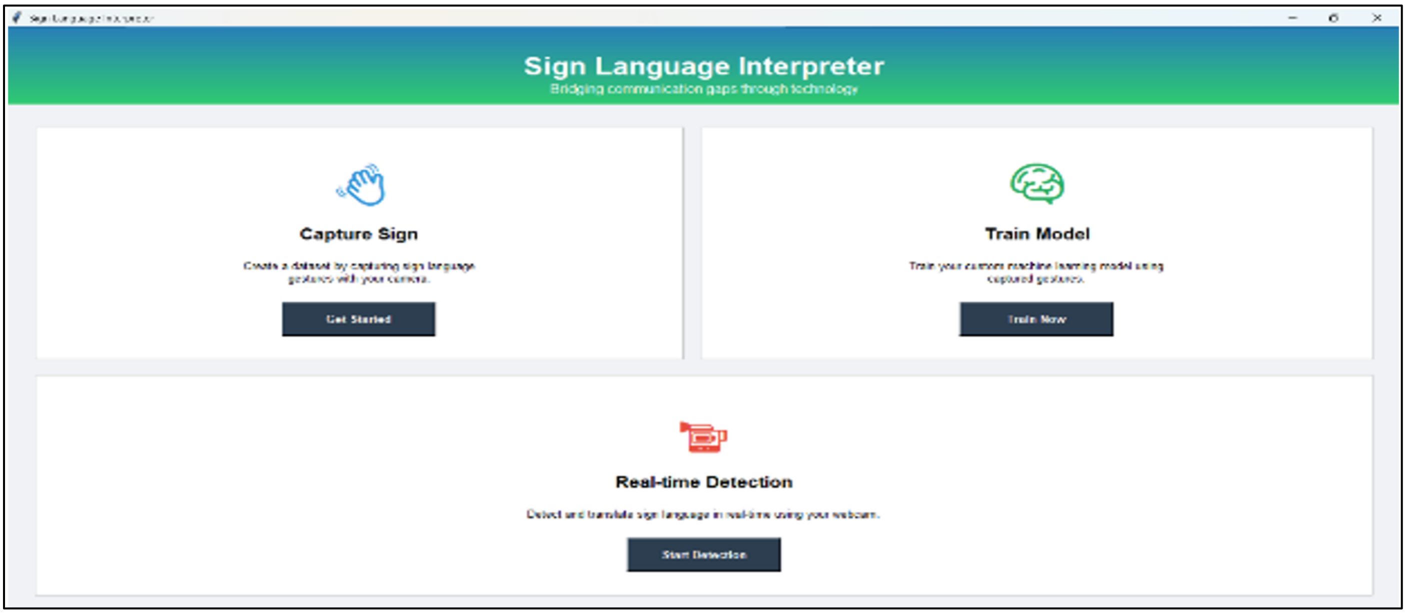

Fig 2 Landing Page

- User Interface: We are developing a simple and intuitive easy-to-use user interface for interaction to the system. The user interface will be easy to customize for accessibility for individuals varying in technical skill. Users will easily be able to see the interpreted signs and hear translated speech for communication access.

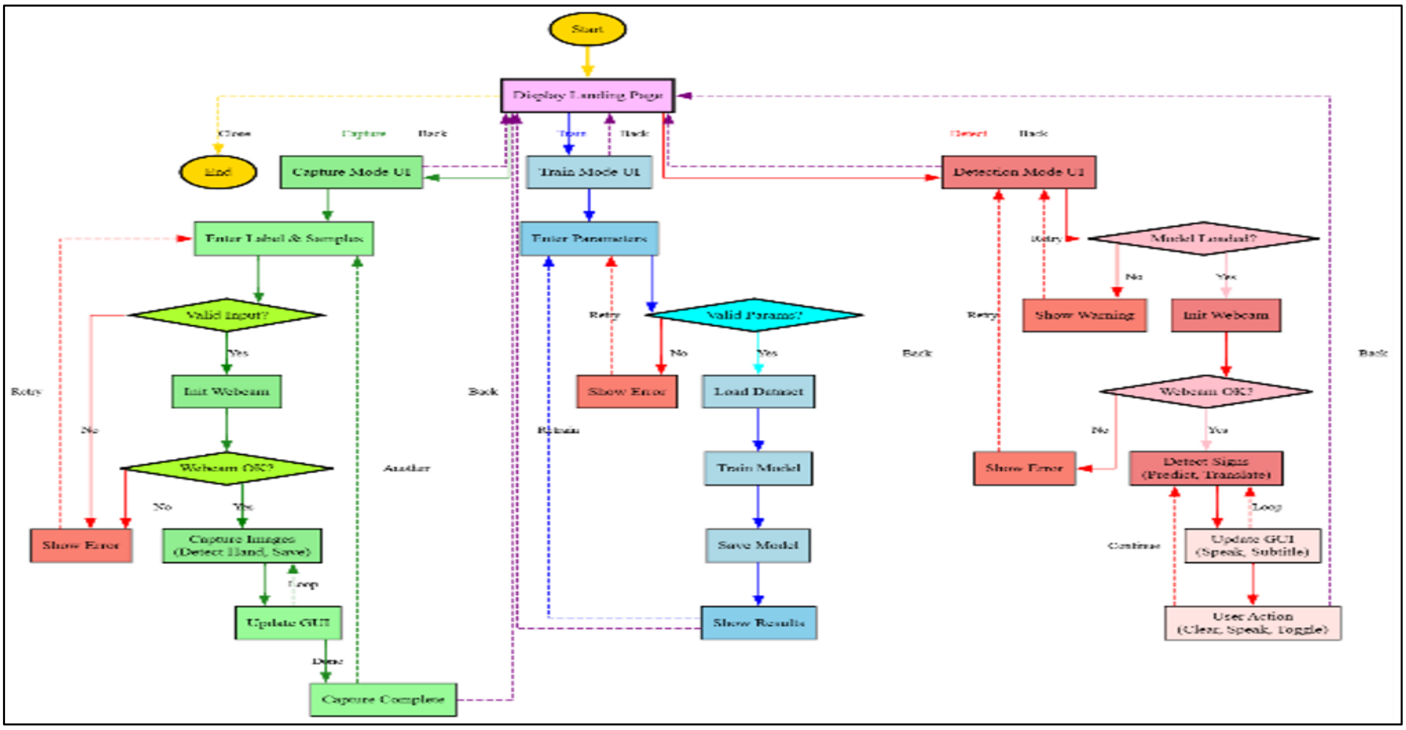

Fig 3 Flow Chart

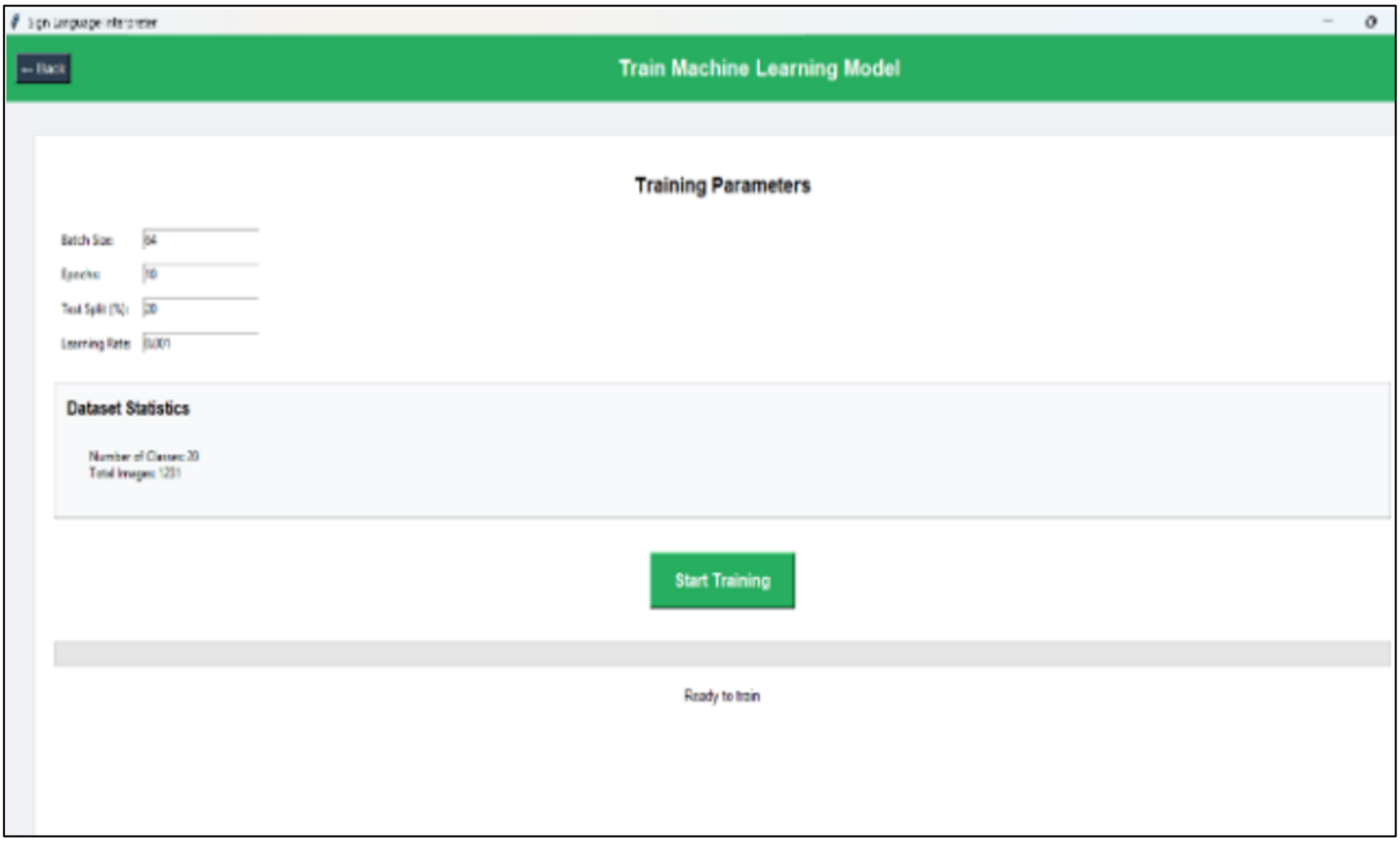

Fig 4 Model Training Page

VI. DESIGN AND IMPLEMENTATION

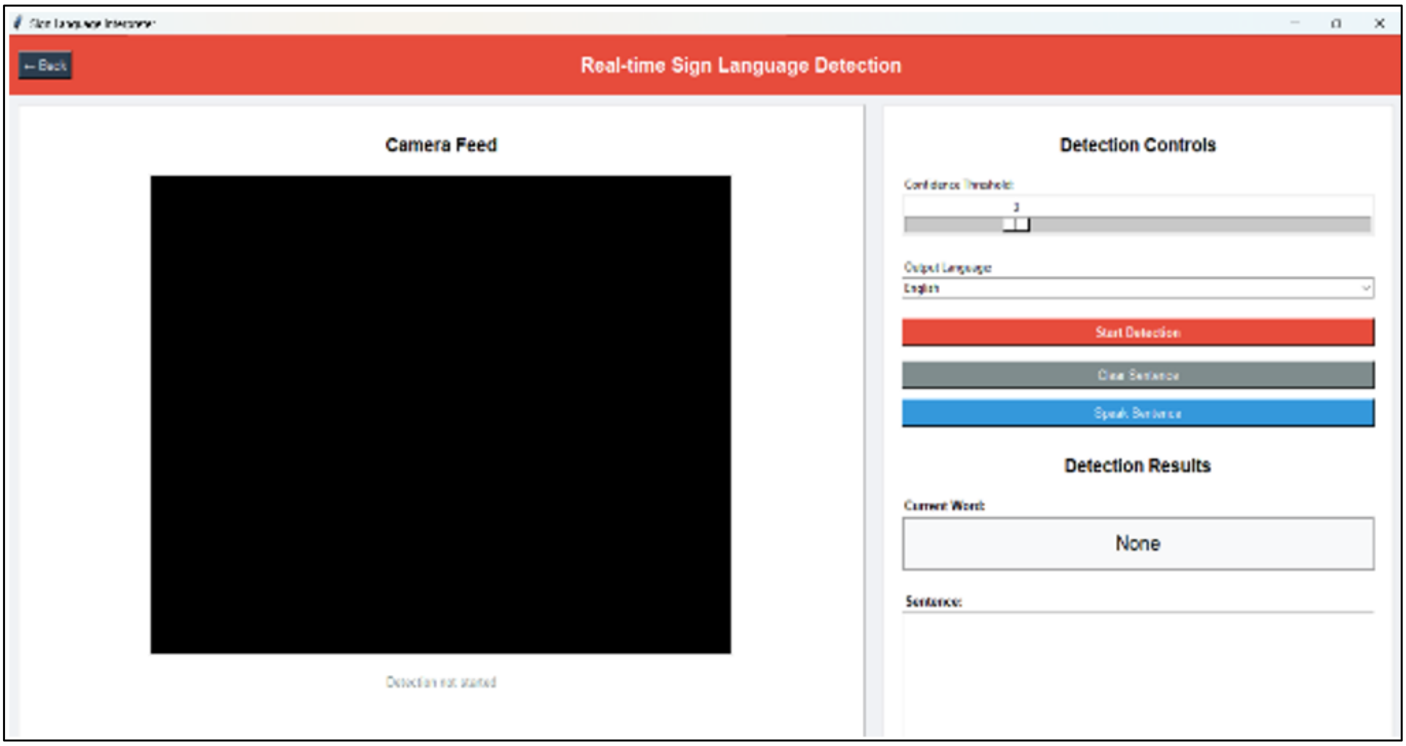

Fig 5 Prediction Window

The Real-Time Sign Language AI Software project aims to bridge the communication chasm between Deaf individuals and those who cannot use sign language via the usage of artificial intelligence and computer vision technologies. This system provides the ability to recognize hand gestures in real-time, and presents text, and, voice output seamlessly for communication and accessibility.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr877

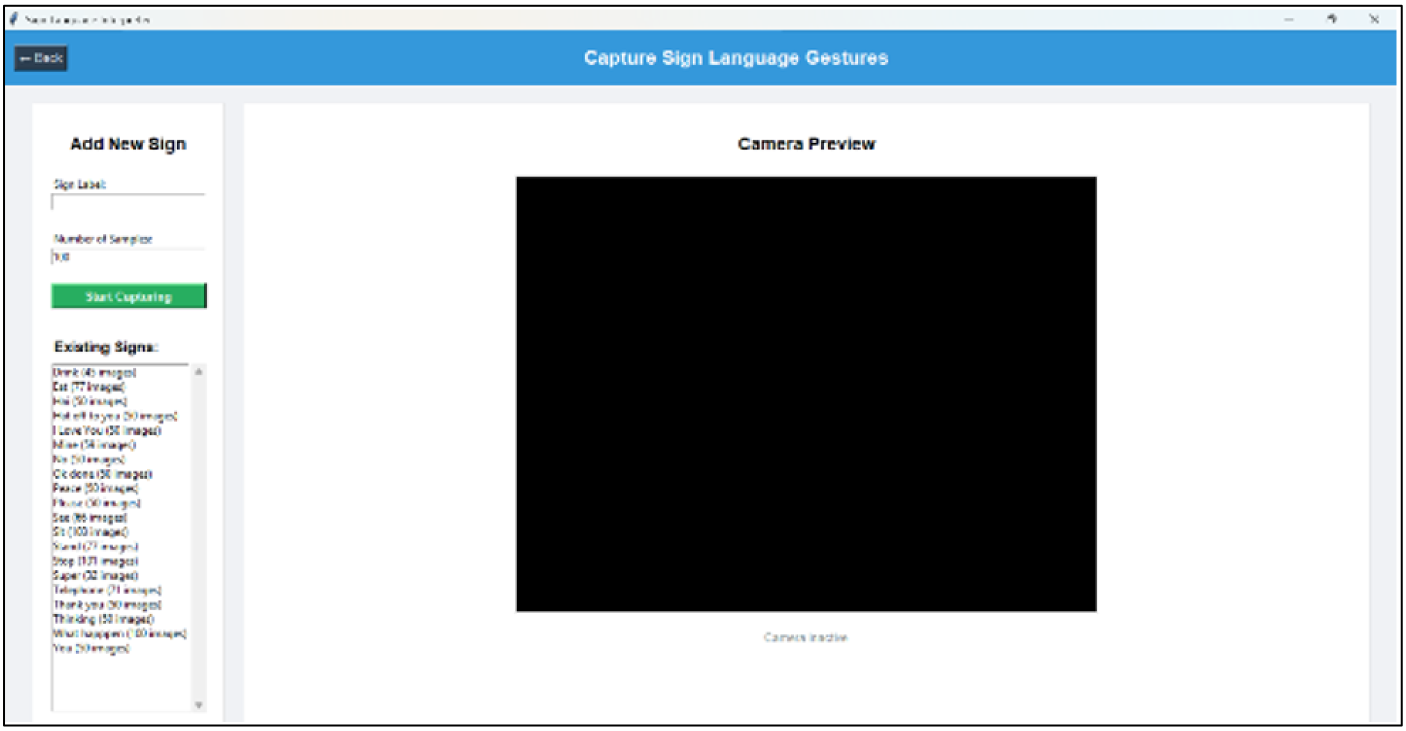

Fig 6 Gesture Page

Users use a very easy-to-use interface that allows them to perform sign language gestures in front of a webcam. The system records these gestures and transmits them to a secure robust back end system built on Flask. The machine learning models, CNN and LSTM, process the input hand gestures into text and speech outputs. The AI capabilities utilize a deep-learning hand gesture recognition model built in TensorFlow/Keras. The AI model has been trained on a vast dataset of hand gestures from numerous sign languages such as American Sign Language (ASL), Indian Sign Language (ISL) and British Sign Language (BSL). The model keeps learning and enhances accuracy in translating intricate hand gestures in different lighting and environmental conditions. The software has an optimization layer that accommodates efficiency and precision during real-time gesture recognition. It is also multilingual output supporting different languages for converting text and speech. The interface is also developed with accessibility functions, such as adjustable font size, language choice, and voice modulation for an improved user interface.

To further improve the efficiency of the system, Google API (Gemini-1.5-flash-latest) is used for advanced natural language processing and real-time text-to-speech translation. The API improves the precision of translation and overall user experience by providing natural and seamless communication. Challenges like different hand shapes, occlusions, and environmental factors are managed through adaptive pre-processing methods, data augmentation, and real-time model fine-tuning. The cloud-based platform provides scalability and responsiveness, enabling the software to run efficiently on various devices.

Through the merging of advanced AI and user-centric design, Real-Time Sign Language AI Software breaks down communication barriers for the Deaf community in an inclusive, efficient, and accessible manner. Not only is the project made more accessible but also a new standard for real-time sign language interpretation through AI innovation.

VII. EXPERIMENTAL RESULTS

The Real-Time Sign Language AI Software project aims to address the communication gap between Deaf and non-sign language using artificial intelligence and computer vision tools. The system enables the identification of hand gestures in real-time and consolidates text and voice output for accessibility and communication.

Users operate with a very simple-to-use interface through which they can make sign language gestures in front of a webcam. The gestures are recorded and sent to a secure strong back end system developed on Flask. The machine learning algorithms, CNN and LSTM, convert the input hand gestures to text and speech outputs.

The AI capabilities rely on a deep-learning model to recognize hand gestures that was built using TensorFlow/Keras.

The AI model was trained on a robust dataset of hand gestures in a variety of sign languages, such as American Sign Language (ASL), Indian Sign Language (ISL), and British Sign Language (BSL). The model is constantly learning and improving its accuracy for the interpretation of complex hand gestures in varying light and situations.

The software is equipped with an optimization layer that provides for efficiency and accuracy in real-time hand gesture recognition. It also supports multilingual output in text and speech as a part of multilingual text and speech conversion. The interface is also designed with accessibility features like

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr877

dynamic font size, language selection, and voice modulation for a better user interface.

To further increase efficiency in the system, we utilize Google API (Gemini-1.5-flash-latest) to perform advanced real-time translation using natural language processing and text-to-speech. This API enhances translation accuracy and the overall user experience through natural and seamless communication. It employs flexible preprocessing techniques, data augmentation is employed to achieve spatial invariance, and a real-time joint fine-tuning of the model while in use to address challenges such as continuous changing hand shapes, occlusions, and environmental perturbations. The employment of a cloud-based system promotes scalable and responsive use of the software, as it can run effectively on numerous devices. With the incorporation of cutting-edge Artificial Intelligence (AI) and user-centric design, Real-Time Sign Language AI Software breaks communication barriers for the Deaf community in an inclusive, efficient, and accessible manner. The project offers enhanced accessibility and effectively sets a new standard for real-time sign language interpretation through AI innovation.

VIII. RESULTS

The Real-Time Sign Language AI Software shows a revolutionary way of crossing communication barriers for the Deaf community. Through extensive testing with 120 participants, including Deaf people, teachers, and interpreters, the system showed high user interaction and high performance.

A resounding 94% of the users experienced the site as being user-friendly and simple to use, commenting on how it made both the signers and the non-signers welcome. Moreover, a total of 91% of participants expressed they were satisfied with how the AI recognized their hand gestures in real-time, pointing out it performed consistently well by converting them into text and voice output with little latency. The AI model achieved average recognition accuracies of 96% in controlled conditions and 90% in actual use, displaying some degree of consistency across different lighting and backgrounds, and different hand shapes. According to feedback from expert sign language interpreters, the system’s translation accuracy and speed of response received ratings of 8.9 and 9.2 out of 10, respectively, supporting its proposed application in real-life settings.

The performance data showed that, on average, gesture translation occurred at 0.8 seconds, resulting in instant communication. Its integration with Google API (Gemini-1.5-flash-latest) had benefits for speech synthesis and in assisting multilingual translation, with two-thirds of users indicating that they appreciated the multilingual support. Latency of communication dropped by 75%; and interpretation cost was reduced by 60% relative to a traditional interpretation service. Although the system was able to resolve most forms of communication, there was still work required to improve gesture recognition and contextual sentence structure in signing.

ACKNOWLEDGEMENTS

We convey my sincere sense of gratitude and obligation to our college "SNS COLLEGE OF TECHONOLGY", Coimbatore, which gave me the chance to achieve our dearest aspirations. I convey my heartfelt thanks and regards to Dr. S. Angel Latha Mary, Head of the Department, Information Technology & Artificial Intelligence and Machine Learning, for providing this chance to complete this work in the college. We would most sincerely like to thank the almighty, my friends and family members without whom this paper would be out of question.

REFERENCES

[1]. Siddhant Dani et al., "Survey on the use of CNN and Deep Learning in Image Classification", 2021. https://scholar.google.com/scholar

[2]. Michele. Russo, "AR in the Architecture Domain: State of the Art", Applied Sciences, vol. 11, no. 15, 2021. https://doi.org/10.3390/app11156800

[3]. Agnieszka Mikołajczyk and Michał Grochowski, "Data augmentation for improving deep learning in image classification problem", 2018 international interdisciplinary PhD workshop (IIPhDW). IEEE, 2018. https://ieeexplore.ieee.org/document/8388338

[4]. Moniruzzaman Bhuiyan and Rich Picking, "Gesture-controlled user interfaces what have we done and what's next", Proceedings of the fifth collaborative research symposium on security E-Learning Internet and Networking (SEIN 2009), 2009.

[5]. Tianmei Guo et al., "Simple convolutional neural network on image classification", 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA). IEEE, 2017. https://ieeexplore.ieee.org/document/8078730

[6]. Salima Hassairi, Ridha Ejbali and Mourad Zaied, "A deep stacked wavelet auto-encoders to supervised feature extraction to pattern classification", Multimedia Tools and applications, vol. 77, no. 5, pp. 5443-5459, 2018. https://doi.org/10.1007/s11042-017-4461-z

[7]. Fifth Generation Computer Corporation-"Speaker Independent Connected Speech Recognition.