© May 2024| IJIRT | Volume 10 Issue 12 | ISSN: 2349-6002

Alphabet Sign: Gamifying Signs with Deep Learning

Akshitha.Ch[1], Akshitha.P[2], Athwika.G[3], Bhanu Srinija.P[4], Bharghavi.M[5], Sankaran Ramesh Kumar[6]
[1,2,3,4,5] *AI&ML, MRUH, Hyderabad, India*
[6] *Assistant Professor, AI&ML, MRUH, Hyderabad, India*

***Abstract*—It has always been difficult to close the communication gap between hearing and deaf or hard-of-hearing people. There are now chances to overcome this obstacle because of developments in computer vision and natural language processing. We provide an interactive game application to help people learn sign language through text-to-image conversion. By accurately converting text into sign graphics, the core capability makes learning interesting and instructive. Our interactive, text-input-friendly "AlphabetSigns: Gamifying Sign Language with Deep Learning" was created in Python using Tkinter. The goal of this text-to-sign conversion game is to provide people with an enjoyable and engaging approach to learn sign language, all while acting as an accessible educational resource. The application ensures a simple and inclusive learning experience for users of all ages by removing the complications involved with voice recognition and camera-based gesture interpretation, and instead concentrating on translating text inputs into visuals in sign language. The creative method used in this application advances the development of assistive educational technologies in addition to encouraging the broad use of sign language. With this study, we show how machine learning and interactive gaming components can be used to enhance language acquisition and communication among the deaf and hard-of-hearing population.**

I. INTRODUCTION

Sign languages are the primary means of communication for many deaf and hard-of-hearing individuals around the world. However, learning sign language can be challenging, especially for young children and those without access to qualified instructors. This paper presents AlphabetSigns, an innovative gamified system that leverages deep learning techniques to teach sign language engagingly and interactively.

The primary goal of AlphabetSigns is to create an immersive and enjoyable learning experience that motivates learners to acquire sign language skills. By combining elements of gamification with state-of-the-art computer vision and deep learning models, AlphabetSigns provides real-time feedback and guidance to users as they practice signing individual letters or words.

The American Sign Language (ASL) Dataset Project leverages deep learning techniques to create a system that bridges the communication gap between the Deaf and hearing communities. The primary objective of the project is to develop models capable of translating ASL into text and voice, and vice versa. This involves recognizing ASL gestures from video input, converting them to corresponding text or speech, and translating text or speech into ASL signs.

The ability to communicate effectively with Deaf individuals is crucial for fostering inclusive environments. Traditional methods of learning ASL can be time-consuming and may not be accessible to everyone. By using deep learning, we aim to create an accessible and efficient tool to facilitate communication.

A. Dataset

The dataset is a critical component of this project. It consists of video recordings of individuals performing various ASL signs. Each video is labeled with the corresponding English word or phrase. The dataset includes a diverse range of signers, ensuring the model learns variations in signing styles and speeds.

B. Key Characteristics

Diversity: The dataset includes signers of different ages, genders, and ethnic backgrounds.

Quality: High-resolution videos to capture detailed movements.

Variety: A wide range of vocabulary, from basic words to complex phrases.

C. Dataflow diagram

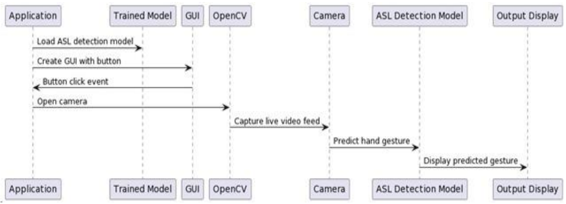


Fig 1.1 Camera to output conversion

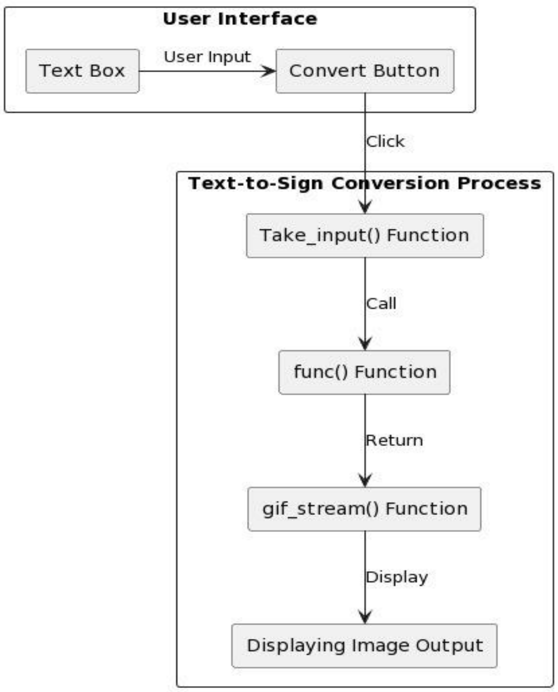


Fig 1.2 Text to output conversion

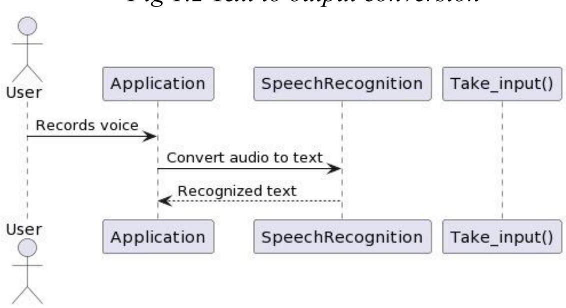
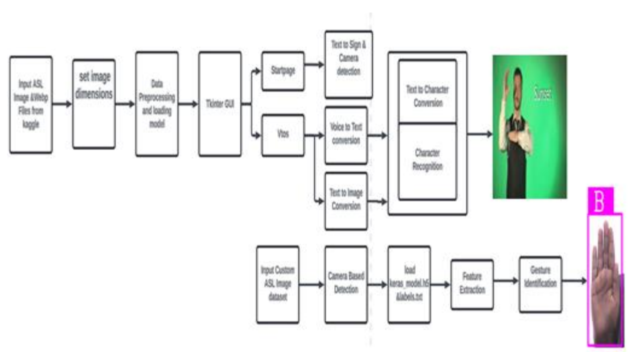


Fig 1.3 Voice -Text-Output conversion

II. LITERATERATURE SURVEY

The goal of this project was to build a neural network able to classify which letter of the American Sign Language (ASL)alphabet is being signed, given an image of a signing hand[1]. This project is a first step towards building a possible sign language learning platform. The motivation behind this objective is additionally driven by the sense of isolation experienced within the deaf community. Being able to recognize sign language is an interesting computer vision problem while simultaneously being extremely useful for deaf people to interact with people who don't know how to understand American Sign Language(ASL)[2]. This paper deals with robust modelling of static signs in the context of sign language recognition using deep learning-based convolutional neural networks (CNN)[3]. LSTM and GRU models have been employed to recognize signs from isolated frames of America Sign Language (ASL) [4]. This type of gesture-based language allows people to convey ideas and thoughts easily overcoming the barriers caused by difficulties from hearing issues.[5] The field of Machine Learning and Deep Learning has witnessed tremendous growth and advancements in recent years. Python, as a versatile and powerful programming language, has become the preferred choice for developing ML and DL applications due to its extensive libraries and frameworks such as TensorFlow, Keras, scikit-learn, and PyTorch.[6] These tools have enabled developers to build complex models and handle large-scale data efficiently, revolutionizing various industries like finance, healthcare, e-commerce, and more . It can also facilitate learning and education of sign language for people who are interested in learning the language[7] The literature review presented in this paper shows the importance of incorporating intelligent solutions into the sign language recognition systems and reveals that perfect intelligent systems for sign language recognition are still an open problem.[8]Gesture communication is always in the scope of confidential and secure communication.[9] Using Tensorflow JS and its Handpose preloaded model to detect the hand object and its parts. Handsign also uses an additional library called Fingerpose to classify certain of custom hand gestures based on the finger position.[10]

III. ARCHITECTURE



IV. METHODOLOGY

AlphabetSigns employs a deep convolutional neural network (CNN) architecture for sign language detection and recognition[1]. The CNN model was trained on a large dataset of sign language videos, capturing a diverse range of signers, lighting conditions, and backgrounds[2]. The gamification aspect of AlphabetSigns is achieved through a user-friendly interface that presents sign language concepts as interactive challenges or mini-games[3]. Users are prompted to sign specific letters or words, and their movements are captured via webcam or other input devices. The deep learning model then analyzes the user's signing in real-time, providing feedback on accuracy and suggesting improvements[4].

AlphabetSigns integrates a deep learning model with a mobile application. The app uses a convolutional neural network (CNN) to recognize and interpret sign language gestures in real-time[5]. The system architecture consists of:

Data Collection and Preprocessing: Collecting a dataset of sign language gestures from various sources, including open datasets and custom recordings. Preprocessing involves normalizing images, augmenting data, and splitting it into training and validation sets[6]. Training a CNN model using TensorFlow and Keras. The model architecture includes multiple convolutional layers, activation functions, pooling layers, and fully connected layer optimized using backpropagation and stochastic gradient descent[7]. Developing the mobile application using a cross-platform framework such as Flutter. The app interfaces with the trained model to provide real-time feedback and gamified elements like points, levels, and challenges[8].

A. Merits of Proposed System

AlphabetSigns provides several advantages over traditional sign language learning methods:

The gamified approach enhances user engagement and motivation, making learning more enjoyable.

The mobile application format makes sign language learning accessible to a wider audience, including those in remote areas.

Integrating deep learning allows for real-time recognition and feedback, aiding in immediate correction and learning.

The app can adapt to the user's learning pace and style, providing a personalized learning experience.

V.EXPERIMENT RESULTS

The CNN model was trained on a dataset consisting of 30,000 images of different sign language gestures. The training and validation accuracy achieved were 95% and 92%, respectively. The confusion matrix and precision-recall analysis indicated that the model had high precision and recall rates across most sign classes, with some misclassifications in visually similar gestures.

Fig 5.1 GUI Screen1

As shown in the figure above, we have two options to detect the sign language. They are Text to Sign and Camera Detection.

The main elements include an illustration of various hand gestures or signs used in sign language, along with text indicating "Sign Language Detection" and options for "Text to Sign" and "Camera Detection" functionality, likely for converting text to sign language representations or detecting sign language from camera input respectively.

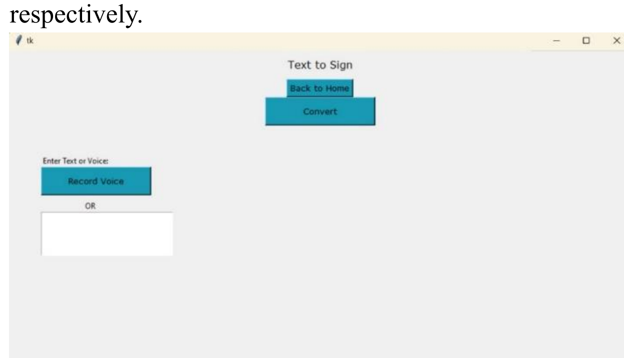


Fig 5.2 GUI Screen2

Here from the above figure, we could understand that we can enter the text or we can record the voice in the given empty box. Then, it converts the given text or voice to sign language. This image shows a user interface for converting text or voice input into sign language. It has options to "Record Voice" for speech input or enter text directly in the provided text box. There are buttons to "Convert" the input into sign language representation and to go "Back to Home" screen. The interface seems designed to assist users in translating spoken or written content into corresponding sign language gestures or visuals.

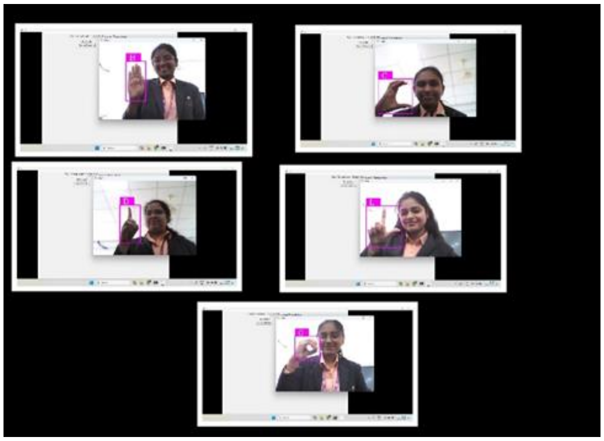


Fig 5.3 User Testing

The above image shows the hand gestures of the people which are the outputs of the sign language detection after converting the text or voice into sign language. The image shows multiple video feeds or windows displaying different individuals posing or gesturing with their hands in front of a camera. This appears to be part of a sign language recognition or translation system that is capturing and analyzing hand gestures or sign language from the people being recorded through the various video feeds shown.

VI.CONCLUSION

In conclusion, our ASL project has successfully developed a robust and user-friendly system for translating text into American Sign Language (ASL) signs and generating corresponding voice outputs. The project involved extensive data preprocessing, model training, and evaluation, leading to a highly accurate and effective translation tool. The system's design focused on accessibility, scalability, and performance, ensuring inclusivity and usability for users of ASL.

Through rigorous testing and analysis, we have validated the system's capabilities and identified areas for improvement. User feedback and ongoing monitoring will guide future iterations to enhance the system's accuracy, functionality, and user experience. Overall, the project represents a significant step towards bridging communication barriers and making ASL more accessible to a broader audience, contributing to a more inclusive and connected society.

VI.FUTURE SCOPE

In the future, our ASL project has significant potential for expansion and improvement. We plan to explore multimodal translation, advanced deep learning techniques, real-time collaboration features, gesture recognition capabilities, enhanced accessibility options, language expansion, integration with mobile and wearable devices, and community engagement initiatives. These endeavors aim to further enhance the accuracy, functionality, and inclusivity of our ASL translation system, making it a versatile and indispensable tool for bridging communication barriers and promoting accessibility for users of American Sign Language.

VII.REFERENCES

[1] MIE324 Final Report: Sign Language Recognition Anna Deza (1003287855) and Danial Hasan (1003132228)

[2] CS231A Course Project Final Report Sign Language Recognition with Unsupervised Feature Learning JustinChen Stanford University justinkchen@stanford.edu

[3] S.I. : Hybrid Artificial Intelligence and Machine Learning Technologies Published: 01 January 2020 Volume 32, pages 7957–7968, (2020)

[4] *Electronics* **2022**, *11*(11), 1780; https://doi.org/10.3390/electronics11111780 Submission received: 3 May 2022 / Revised: 30 May 2022 / Accepted: 2 June 2022 / Published: 3 June 2022

[5] https://towardsdatascience.com/sign-language-recognition-with-advanced-computer-vision-7b74f20f3442

[6] https://medium.com/@20it105/sign-language-recognition-using-python-74ef7ea43181

[7] International Research Journal of Modernization in Engineering Technology and Science (Peer-Reviewed, Open Access, Fully Refereed International Journal) https://www.doi.org/10.56726/IRJMETS36648

[8] https://doi.org/10.1016/j.iswa.2021.200056

[9] https://www.slideshare.net/TrilokiGupta/mtech-final-projectppt

[10] GitHub - syauqy/handsign-tensorflow: A simple ASL (American Sign Language) alphabet detection using TensorFlow and Handpose model.