



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 14, Issue 2, February 2025**

# **SignEase: Intuitive Sign Language Translation with Enhanced Features**

**Renuka Pange<sup>1</sup>, Shivam Shevkari<sup>2</sup>, Ishika Bhartiya<sup>3</sup>, Suraj Chavan<sup>4</sup>, Prof. Pooja Pawar<sup>5</sup>**

Department of Computer Engineering, Suman Ramesh Tulsiani Technical Campus, Faculty of Engineering, Kamshet,  
Pune, India<sup>1,2,3,4</sup>

Department of Computer Engineering, Suman Ramesh Tulsiani Technical Campus, Faculty of Engineering, Kamshet,  
Pune, India<sup>5</sup>

**ABSTRACT:** Communication is a crucial element of human interaction, but unfortunately, millions of people with hearing and speech impairments encounter obstacles in their daily lives. Approximately 430 million individuals worldwide experience disabling hearing loss, with a significant portion relying on sign language as their primary mode of communication. Unfortunately, there is a substantial gap in communication between individuals who use sign language and those who rely on spoken language. This paper introduces SignEase, a novel platform that aims to bridge the gap between sign and spoken language by utilizing machine learning and MediaPipe's landmark detection system for precise sign language recognition. The platform integrates cutting-edge text-to-speech (TTS) technology, which converts recognized signs into spoken language, improving accessibility and promoting inclusivity. The sign language recognition system utilizes a CNN model that has been trained on a dataset of hand landmarks extracted using MediaPipe, allowing for real-time translation of sign language into both text and speech. By combining advanced technology with user-focused features, SignEase strives to enhance communication between the deaf and hearing communities, ultimately promoting social inclusion, educational opportunities, and an improved quality of life. The article examines the approach, system design, and real-time efficiency of the platform, demonstrating its potential as a ground-breaking solution for addressing communication challenges experienced by individuals with hearing impairments.

## **I. INTRODUCTION**

Communication is a fundamental aspect of human interaction, yet millions face barriers due to hearing and speech impairments. Over 5% of the world's population, approximately 430 million people, experience disabling hearing loss, a figure projected to surpass 700 million by 2050. The impact of hearing loss is particularly pronounced in low and middle-income countries, where 80% of affected individuals reside. Among those with profound hearing loss, many rely on sign language as their primary means of communication, with over 300 distinct sign languages spoken by an estimated 72 million people globally.

Despite the widespread use of sign language, a significant communication gap exists between sign language users and those who rely on spoken languages. This disconnect poses daily challenges, limiting access to education, employment opportunities, and social inclusion for the deaf and hard-of-hearing communities. Addressing this disparity is vital for fostering inclusivity and empowering individuals with hearing and speech impairments.

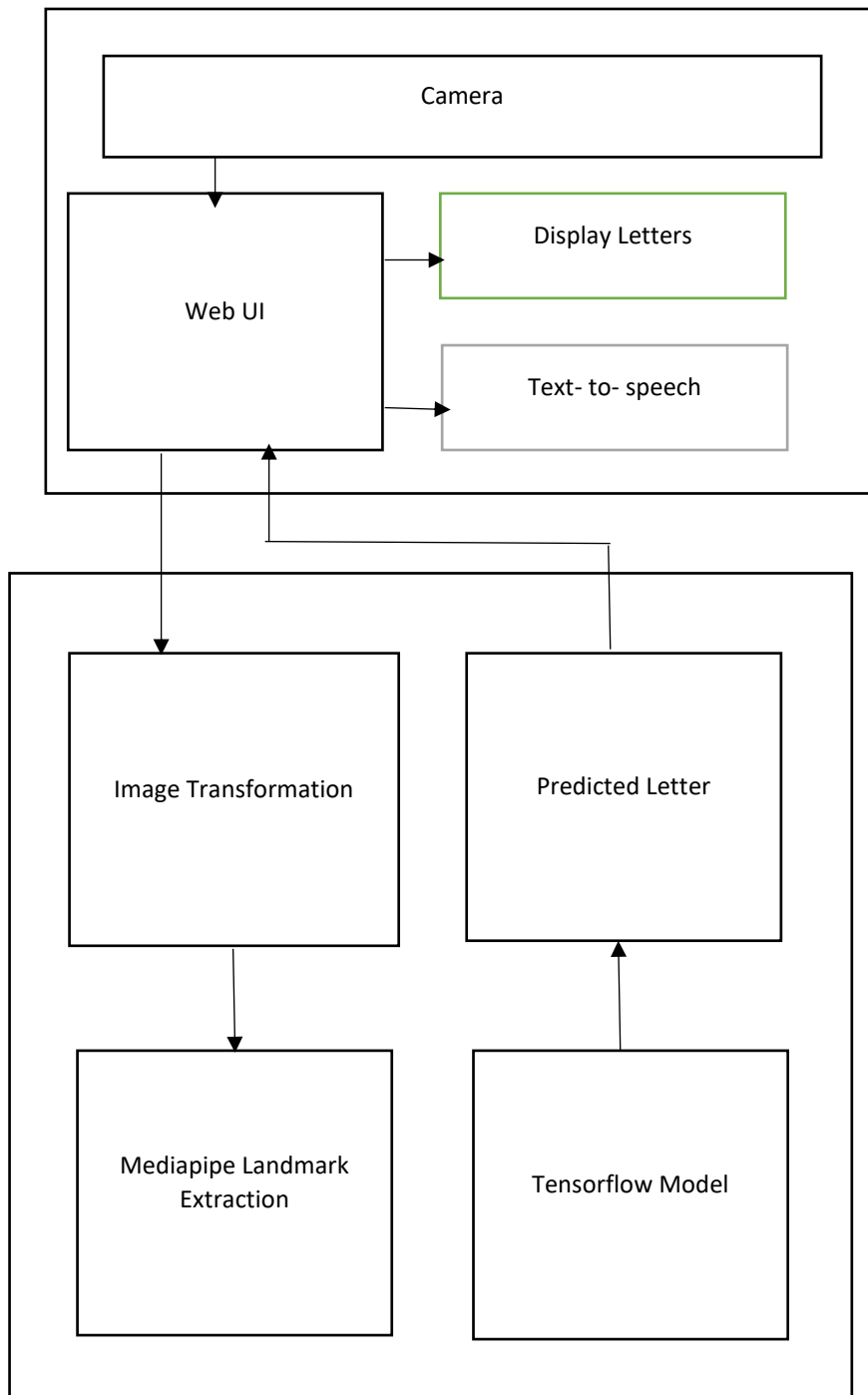
This paper introduces SignEase, a novel platform designed to bridge the communication gap between sign language users and spoken language users. Leveraging the power of machine learning and MediaPipe's landmark detection system, SignEase facilitates accurate recognition of sign language alphabets. Beyond recognition, the platform enhances accessibility by integrating text-to-speech (TTS) translation capabilities, enabling seamless conversion of recognized signs into spoken language using advanced TTS engines such as Google TTS, IBM Watson, and Azure Speech. By combining cutting-edge technology with user-centric features, SignEase aims to create a more inclusive and accessible communication experience for all.

## II. LITERATURE SURVEY

Recent advancements in machine learning, particularly through the TensorFlow library, have enabled the development of sophisticated web-based applications like SignEase. These applications facilitate real-time sign language recognition and translation. SignEase also incorporates Text-to-Speech and multilingual support, enhancing communication and accessibility for sign language users.

1. In an article by Wilson and Anspach [1], they applied neural networks to sign language translation, concentrating on extracting significant characteristics (such as motion, location, and configuration) from video images. The authors developed a modular hybrid approach using multiple methodologies (including neural networks) to mediate communication between deaf signers and hearing individuals. The methodology utilized a trained neural network to classify handshapes of American Sign Language using Fourier descriptors as input. The authors demonstrated that the trained network was invariant with camera type and segmentation methods.
2. In a recent study by Jiang et al. [2], a new sign language retrieval framework, Semantically Enhanced Dual-Stream Encoder (SEDS), was developed to address some of the limitations of traditional RGB-based sign language video representations. Within this design, pose and RGB modalities are used as a paired way of capturing both local and global information from sign language videos. A Cross Gloss Attention Fusion (CGAF) module enables the context-aware fusion of the video modalities, along with a Pose-RGB Fine-Grained Matching Objective to enhance the representation of the pooled features. The network is designed to be lightweight for end-to-end training, and extensive reports demonstrate that SEDS outperforms competition across datasets proposed in the field.
3. Another study by Harini et al. [3] focuses on developing a sign language translation system for real-time communication aids for bridging communication barriers between the hard-of-hearing population and the general public. The sign language translation system is made up of four modules: image capturing, image preprocessing, gesture classification, and predicting the sign gestures. Using OpenCV and Python, the image capturing and preprocessing modules capture sign gestures, resize them, convert them to grayscale, and download noise to improve prediction accuracy. A Convolutional Neural Network (CNN) is used for gesture classification and predicting modules. Their work demonstrates the effective development of a computer vision technique to translate sign language in real time and offer a possible communication aid for the deaf community.
4. Ananthanarayana et al. [4] published thorough research on deep learning-based approaches to sign language translation, specifically encoding sign language input without gloss-based information. The authors applied three types of machine translation models including sequence-to-sequence, attention-based networks, and reinforcement learning, as well as a transformer model on German, American, and Chinese Sign Language. The authors concluded that a transformer model with ResNet50 embeddings or pose-based landmarks achieved the greatest BLEU2-BLEU4 scores across all datasets, particularly on the German Sign Language dataset. This paper demonstrates the advantages of advanced neural translation models applied to continuous sign language translation in differing languages.
5. He [6] proposed a sign language translation system that employs deep learning to improve recognition accuracy via state-of-the-art neural network developments. The translation system uses first Faster R-CNN to locate hands, followed by a 3D CNN to extract features, and finally, an LSTM-based encoder-decoder network to recognize sign language sequences. The integrated approach works well to exploit the contextual nature of sign language, achieving a recognition rate as high as 99% on basic vocabulary datasets. This work illustrates the benefits of deep learning methods to complete the task of sign language recognition in practical settings.
6. Orbay and Akarun [7] explored neural sign language translation using a novel tokenization scheme to reduce the dependence on expensive, annotated data. Their method presents novel semi-supervised tokenization methods via adversarial, multitasking, and transfer learning. Their extensive experiments showed that with only sentence-level annotation their methods raised other reports of state-of-the-art results on translation tasks achieving task performance of 13.25 (BLUE-4) and 36.28 (ROUGE); note BLUE-4 and ROUGE scores improved by 4 and 5 points respectively over previous developments. ORBAY did useful work to move the field forward by improving translation outcomes while having little data annotation.

7. Bragg et al. [8] give an interdisciplinary perspective on sign language recognition, generation, and translation arguing expertise in a variety of topics across computer vision, natural language processing, human-computer interaction, and linguistics is needed to be successful in solving the problem of sign language translation. Their work is based on a workshop to disseminate pieces of work from the workshop with shared expertise and claims to address some key challenges and field-level needs. The document provides an overview of potential work related to technologies already accepted as state-of-the-art and provides an overview of areas that require more attention and calls to action. Their work aims to advance the body of work in sign language work with consideration of collaborative knowledge.
8. Oliveira et al. [9] discussed the Virtual Sign platform, a tool designed to translate sign language to/from text, improving communication for hearing-impaired university students. Developed in 2015, it addresses challenges in sentence structure comprehension and the lack of available sign language interpreters. Positive feedback was received from tests and a pilot experiment. The platform is being further developed with input from sign interpreters across Europe, and new functionalities are continuously being added to enhance its educational support for hearing-impaired students.
9. Parton [10] consolidated and detailed a broad set of literature regarding sign language recognition and translation using a variety of artificial intelligence approaches. The paper discusses earlier projects "Ralph" and CyberGloves, using robotics and VR sensors, and even the CopyCat game and other camera-based systems designed for accessing and learning sign language interactively. Parton also detailed sign-recognition software that has used Hidden Markov Models and neural networks and 3D avatars like "Tessa". Parton highlighted work on natural language processing such as Poland's "PHOTOS" and the intelligent tutoring system being developed in the ICICLE project as work to support deaf students. The review helps to inform readers of both theory and applications for educational contexts while highlighting the interdisciplinary nature of the technologies and applications.
10. Núñez-Marcos et al. [11] furnish a comprehensive overview of sign language machine translation (SLMT), examining classic and neural translation systems in detail, with a specific focus on translating video instances of sign language to spoken language, and the reverse, highlighting the significance of SLMT in reducing communication barriers between deaf and hearing people. Their paper also provides an overview of various datasets used in this arena, for instance, the RWTHPHOENIX-2014T data, and includes a discussion of challenges and future work in this field. Their survey emphasizes the need to develop more effective translation automation to facilitate the accessibility of communication.
11. Kumar et al. [15] implemented a sign language translation system that utilizes visual cues and deep learning techniques to address the limitations of traditional glove-based systems. Their solution included the use of Computer Vision for American Sign Language (ASL) gloss recognition, and Neural Machine Translation (NMT) to convert sign language formatted video instances to English encoded sentences. The end-to-end neural network-based system showed competency in the recognition of discrete ASL signs as well as continuous inputs of sign language as a medium of communication with non-signing people. Ultimately, their implementation signifies a useful application of an automated translation platform designed for practical in-situ use.
12. Sonare et al. [19] devised a video-based sign language interpreting system that leverages machine learning to facilitate communication for deaf and mute individuals. Their system consists of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to recognize and translate continuous hand gestures for sign language with 92.4% accuracy. The system converts a video into a written sign language translation to text and then text to speech using Text-To-Speech API, which is open-source. Their system provides a potential for communication if automatic sign language recognition and translation into speech occurs in real-time for various purposes, such as presentations and video links.

**III. SYSTEM ARCHITECTURE****Figure 3: System Architecture**

This diagram shows the flow of a sign language recognition system, starting with a camera that captures hand sign images. The images are processed through a web-based interface, where they are transformed, analysed using MediaPipe for landmark extraction, and passed to a TensorFlow model for prediction. The recognized letters are displayed and converted into audio using text-to-speech for accessible communication.

#### IV. METHODOLOGY

Sign Ease is a machine learning model that is developed to accurately translate sign language into text and speech in real-time using Machine Learning.

##### Model Building Process:

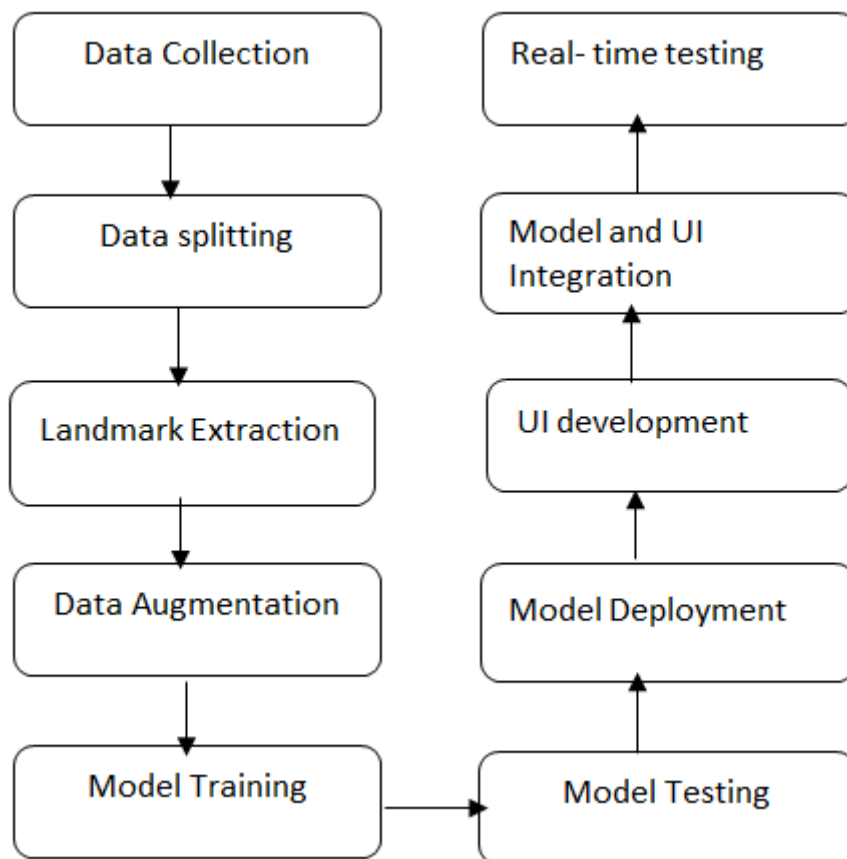


Figure 1: Step-by-step process for building the prediction model

The figure depicts the process of developing and testing a machine learning model, from data collection to model training, UI development, implementation, and real-time testing.

The above flowchart outlines the process of developing and testing a machine learning model, consisting of the following steps:

1. Data Collection: Gathered 168,000 images with 6,000 per alphabet, including two special signs for 'space' and 'delete.' Images captured under varying lighting and distances from the camera.
2. Data Splitting: The dataset is divided into 80% for training and 20% for testing to ensure a balanced evaluation.
3. Landmark Extraction: Utilized MediaPipe to extract 21 hand landmarks in the x, y, and z axes for each image. Key points accurately mapped to hand geometry.
4. Data Augmentation: Applied flipping, rotation, and scaling to augment landmark coordinates. Normalized coordinates with a bounding box to handle varying hand positions.
5. Model Training: Trained a CNN model using extracted hand landmarks to classify sign language symbols.
6. Model Testing: Evaluated model performance on the testing dataset to ensure generalization.
7. Model Deployment: Integrated the trained model into a Python app, enabling real-time image capture via webcam.

8. UI Development: Developed a responsive and interactive web interface using React for frontend functionality.
9. Model & UI Integration: Backend powered by Flask API hosting the model, seamlessly integrated with React frontend.
10. Real-Time Testing: Conducted end-to-end testing of the system to validate its real-time performance.

### V. EXPERIMENTAL RESULTS

The SignEase system was meticulously developed to accurately translate sign language into speech, serving as an innovative tool to bridge the communication gap for individuals with hearing impairments. This section outlines the results and evaluation of the system, highlighting its effectiveness and robustness under various conditions.

#### Model Evaluation:

The model's performance was evaluated using multiple parameters to ensure its accuracy and reliability. When tested against the designated test dataset, the model achieved an outstanding accuracy of 99.6%, indicating its high level of precision in sign language detection and classification. Additionally, the Macro F1-score and Weighted F1-score were calculated to provide deeper insights into the model's balanced performance across all classes. The model attained a Macro F1-score of 0.9286, reflecting its ability to perform well across all classes without being biased toward any particular class, and a Weighted F1-score of 0.9608, showcasing its effectiveness in handling imbalanced datasets. To further analyse the model's performance, a confusion matrix was generated, which provides a comprehensive breakdown of its predictions. This matrix highlights key evaluation metrics such as precision, recall, and overall classification accuracy for each sign language class.

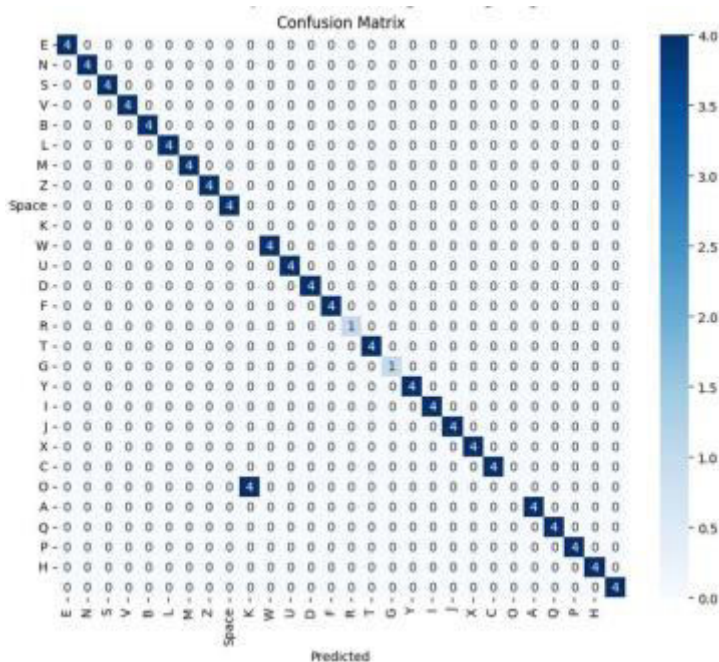


Figure 4A: Confusion Matrix

This figure demonstrates the model accuracy evaluated using the confusion matrix.

The high values observed in the diagonal entries of the matrix indicate that the system effectively distinguishes between the various sign language gestures, minimizing errors.

**Real-World Testing:** In addition to dataset-based evaluations, the SignEase system was subjected to real-world testing to verify its practical applicability. The application was tested in diverse environmental conditions, including variations

in lighting, backgrounds, and hand orientations. These tests demonstrated that the system could effectively detect and translate sign language gestures in real-time scenarios, reinforcing its robustness and usability.

The following images illustrate the system's performance during real-world testing:

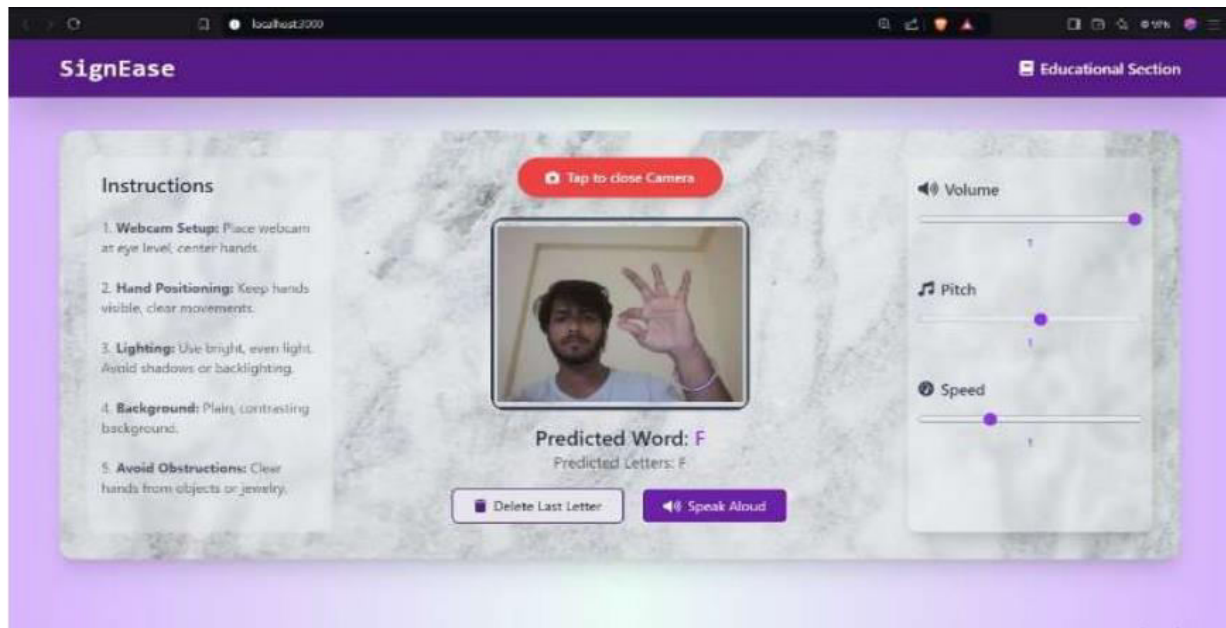


FIGURE 4B: REAL-TIME SIGN DETECTION

This figure shows the user performing the letter 'F' using a sign language gesture, which is accurately predicted and displayed by the system.

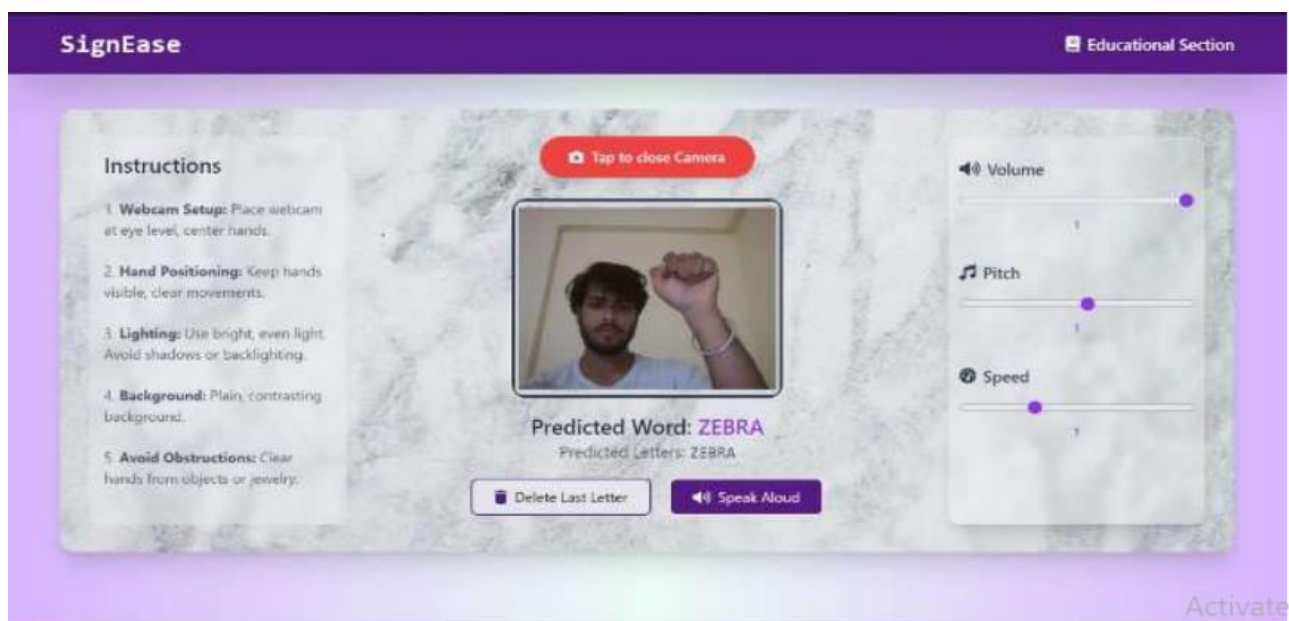


Figure 4C: Word Formation



This figure illustrates how the system displays words by predicting individual letters, with the added capability of providing audio output for the complete word.

## VI. CONCLUSION

SignEase is a web application designed to bridge the communication gap between sign language users and non-signers, fostering inclusivity and accessibility. By integrating machine learning with real-time image processing and text-to-speech translation, SignEase effectively translates sign language alphabets and provides multilingual support, enabling Deaf and hard-of-hearing individuals to communicate more easily. Additionally, the platform allows users to customize voice settings for a more personalized experience.

Beyond serving as a communication tool, SignEase also functions as an educational resource, helping users learn and practice the sign language alphabet. The application is built to be inclusive, catering not only to individuals with disabilities but to everyone. Leveraging TensorFlow for precise sign language detection and utilizing Google TTS, IBM Watson, and Azure Speech for text-to-speech, SignEase aims to transform real-time communication.

While SignEase is not intended to replace professional interpreters, it proves to be a valuable tool in informal and educational contexts. The development team envisions expanding the platform's features and language support, further enhancing its role as an inclusive tool for global communication. Despite its exceptional performance, certain limitations were identified during testing, such as gesture overlap and environmental variability, which offer opportunities for further refinement. By addressing these limitations, SignEase can continue to evolve and meet the demands of real-world applications, ensuring inclusivity across varied use cases.

## REFERENCES

- [1] Beth J. Wilson and Gretel Anspach, "Neural networks for sign language translation," Proc. SPIE 1965, Applications of Artificial Neural Networks IV, (2 September 1993). [<https://doi.org/10.1117/12.152560>](<https://doi.org/10.1117/12.152560>)
- [2] Longtao Jiang, et al., "SEDS: Semantically Enhanced Dual-Stream Encoder for Sign Language Retrieval." arXiv preprint, arXiv:2407.16394(2024). <https://arxiv.org/abs/2407.16394>](<https://arxiv.org/abs/2407.16394>)
- [3] R. Harini, R. Janani, S. Keerthana, S. Madhubala, and S. Venkatasubramanian, "Sign Language Translation," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020, pp. 883-886. [<https://doi.org/10.1109/ICACCS48705.2020.9074370>](<https://doi.org/10.1109/ICACCS48705.2020.9074370>)
- [4] Tejaswini Ananthanarayana, Priyanshu Srivastava, Akash Chintha, Akhil Santha, Brian Landy, Joseph Panaro, Andre Webster, Nikunj Kotecha, Shagan Sah, Thomastine Sarchet, Raymond Ptucha, and Ifeoma Nwogu, "Deep Learning Methods for Sign Language Translation," ACM Trans. Access. Comput. 14, 4, Article 22 (December 2021), 30 pages. [<https://doi.org/10.1145/3477498>](<https://doi.org/10.1145/3477498>)
- [5] Xiujuan Chai, Guang Li, Yushun Lin, Zhihao Xu, Yili Tang, and Xilin Chen, "Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS Beijing, China". [[vipl.ict.ac.cn](mailto:xiujuan.chai@vipl.ict.ac.cn)](<mailto:xiujuan.chai@vipl.ict.ac.cn>)
- [6] S. He, "Research of a Sign Language Translation System Based on Deep Learning," 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), Dublin, Ireland, 2019, pp. 392-396. [<https://doi.org/10.1109/AIAM48774.2019.00083>] (<https://doi.org/10.1109/AIAM48774.2019.00083>)
- [7] A. Orbay and L. Akarun, "Neural Sign Language Translation by Learning Tokenization," 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), Buenos Aires, Argentina, 2020, pp. 222-228. [<https://doi.org/10.1109/FG47880.2020.00002>](<https://doi.org/10.1109/FG47880.2020.00002>)
- [8] Danielle Bragg, Oscar Koller, Mary Bellard, Larwan Berke, Patrick Boudreault, Annelies Braffort, Naomi Caselli, Matt Huenerfauth, Hernisa Kacorri, Tessa Verhoef, Christian Vogler, and Meredith Ringel Morris, "Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective," Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19). [<https://doi.org/10.1145/3308561.3353774>](<https://doi.org/10.1145/3308561.3353774>)

- [9] T. Oliveira, P. Escudeiro, N. Escudeiro, E. Rocha, and F. M. Barbosa, "Automatic Sign Language Translation to Improve Communication," 2019 IEEE Global Engineering Education Conference (EDUCON), Dubai, United Arab Emirates, 2019, pp. 937-942. [<https://doi.org/10.1109/EDUCON.2019.8725244>](<https://doi.org/10.1109/EDUCON.2019.8725244>)
- [10] Becky Sue Parton, "Sign Language Recognition and Translation: A Multidisciplined Approach From the Field of Artificial Intelligence," The Journal of Deaf Studies and Deaf Education, Volume 11, Issue 1, Winter 2006, Pages 94-101. [<https://doi.org/10.1093/deafed/enj003>](<https://doi.org/10.1093/deafed/enj003>)
- [11] Adrián Núñez-Marcos, Olatz Perez-de-Viñaspre, Gorka Labaka, "A Survey on Sign Language Machine Translation," Expert Systems with Applications, Volume 213, Part B, 2023, 118993. [<https://doi.org/10.1016/j.eswa.2022.118993>](<https://doi.org/10.1016/j.eswa.2022.118993>)
- [12] "Sign Language Translation Using Deep Convolutional Neural Networks," KSII Transactions on Internet and Information Systems, Korean Society for Internet Information (KSII), 2020. [<https://doi.org/10.3837/tiis.2020.02.009>](<https://doi.org/10.3837/tiis.2020.02.009>)
- [13] Y. Madhuri, G. Anitha, and M. Anburajan, "Vision-based Sign Language Translation Device," 2013 International Conference on Information Communication and Embedded Systems (ICICES), Chennai, India, 2013, pp. 565-568. [<https://doi.org/10.1109/ICICES.2013.6508395>](<https://doi.org/10.1109/ICICES.2013.6508395>)
- [14] U. Farooq, M.S.M. Rahim, N. Sabir, et al., "Advances in Machine Translation for Sign Language: Approaches, Limitations, and Challenges," Neural Comput & Applic, 33, 14357-14399 (2021). [<https://doi.org/10.1007/s00521-021-06079-3>](<https://doi.org/10.1007/s00521-021-06079-3>)
- [15] S. S. Kumar, T. Wangyal, V. Saboo, and R. Srinath, "Time Series Neural Networks for Real-Time Sign Language Translation," 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 2018, pp. 243-248. [<https://doi.org/10.1109/ICMLA.2018.00043>](<https://doi.org/10.1109/ICMLA.2018.00043>)
- [16] Mehrez Boulares and Mohamed Jemni, "Mobile Sign Language Translation System for Deaf Community," Proceedings of the International Cross-Disciplinary Conference on Web Accessibility (W4A '12). [<https://doi.org/10.1145/2207016.2207049>](<https://doi.org/10.1145/2207016.2207049>)
- [17] Floris Roelofsen, Lyke Esselink, Shani Mende-Gillings, and Anika Smeijers, "Sign Language Translation in a Healthcare Setting," Proceedings of the Translation and Interpreting Technology Online Conference, 2021, pages 110-124. INCOMA Ltd.
- [18] D. Jurafsky and J. H. Martin, "Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (Part 1)." [<https://www.thuvienso.net/upload/library/204335/preview/ebookspeech-and-language-processing-an-introduction-to-natural-languageprocessing-computational-linguistics-and-speech-recognition-part-1-danieljur.pdf>](<https://www.thuvienso.net/upload/library/204335/preview/ebookspeech-and-language-processing-an-introduction-to-natural-languageprocessing-computational-linguistics-and-speech-recognition-part-1-danieljur.pdf>)
- [19] B. Sonare, A. Padgal, Y. Gaikwad, and A. Patil, "Video-Based Sign Language Translation System Using Machine Learning," 2021 2nd International Conference for Emerging Technology (INCET), Belagavi, India, 2021, pp. 1-4. [<https://doi.org/10.1109/INCET51464.2021.9456176>](<https://doi.org/10.1109/INCET51464.2021.9456176>)
- [20] H. V. Verma, E. Aggarwal, and S. Chandra, "Gesture Recognition Using Kinect for Sign Language Translation," 2013 IEEE Second International Conference on Image Information Processing (ICIIP2013), Shimla, India, 2013, pp. 96-100. [<https://doi.org/10.1109/ICIIP.2013.6707563>](<https://doi.org/10.1109/ICIIP.2013.6707563>)
- [21] Estrada Jiménez, L.A., Benalcázar, M.E., and Sotomayor, N., "Gesture Recognition and Machine Learning Applied to Sign Language Translation," VII Latin American Congress on Biomedical



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details