International Journal of Computational Science and Engineering Research
ISSN: XXXX- XXXX(Online) , http://www.ijcser.com/
Regular Issue , Vol. 2, Issue. 2 (April - June), 2025 , Pages: 15 - 21
Received: 24 February 2025 ; Accepted: 12 April 2025 ; Published: 10 June 2025
Original Paper : Action Research - Problem-Solution Research Paper
https://doi.org/10.63328/IJCSER-V2I2P3

Empowering Communication: A Web Application for Deaf, Mute, and Sign Language Interpretation

Konda Hari Krishna [1*], B. Suchitha [2], B. Pravallika [3], E. Deepika [4], G. Jaahnavi [5]

[1*] Department of Computer Science and Engineering, Mohan Babu University, Tirupathi, AP, India ; khk396@gmail.com,
[2,3,4,5] Department of Computer Applications, Mohan Babu University, Tirupathi, AP, India

* *Corresponding Author:* khk396@gmail.com

Abstract: The inability to communicate effectively because of communication barriers severely restricts deaf and mute individuals from getting in touch with those who do not understand sign language. A new web-based system has been developed to provide an effortless communication solution between users who are deaf, mute and sign language users. The system implements Django as its backend framework together with text-to-sign language conversion and speech-to-text capabilities to establish successful communication. Through the application users have the option to provide text or audio content that gets translated into American Sign Language (ASL) animations through hand gesture visualization from MediaPipe. The system enables automatic speech recognition through the SpeechRecognition library and functions as an ASR tool to convert spoken words into text. The Google Translator API enables native language translation of this recorded text which expands communication possibilities to different user groups. From datasets obtained through Kaggle the training of a Convolutional Neural Network (CNN) model reaches 99% accuracy for sign recognition leading to accurate communication. The application creates a user-friendly interface which enables people with speech and hearing difficulties to get real-time sign language animations as visual outputs. The novel system surpasses typical assistive communication solutions because it cuts out the requirement for human sign language interpreters to provide inclusive communication channels. This proposed web application will increase accessible interactions thus promoting more social integration opportunities for both deaf and mute users. The system will be improved through language expansion of sign language interpretation and a mobile app adaptation for maximized accessibility.The system incorporates components related to Communication accessibility, sign language recognition, speech-to-text, ASL animation and employs CNN models in conjunction with Django framework,

Keywords: Media Pipe, Automatic Speech Recognition, Google Translator API , Web-Based Assistive.

1. Introduction

Communication plays a vital role in human social interactions even though millions of people across the world experience communication difficulties from hearing or speech disabilities. Traditional sign language serves as the main communication method for deaf and mute people though it provides insufficient communication bridge when interacting with people who don't recognize sign language. A novel web application with machine learning capabilities and natural language processing technology and computer vision components aids effortless communication among users who are deaf or mute or communicate without sign language. The developed system applies Django as its base framework because it offers a secure platform for building scalable applications. Through these features the framework enables users to become registered members with personal account management. The system core capability enables text and audio conversion into hand movements through the MediaPipe framework which provides top-tier pose detection and hand position tracking functions. The system develops a conceptual link between verbal and visual messaging that creates user-friendly conversations.

The core operation of the system depends on the speech-to- text module that uses automatic speech recognition (ASR) approaches to convert spoken words into written text. The text processing through Google Translator API enables language conversion to different Indian regional languages thereby broadening access for

Vol. 2, Issue 2, 2025, https://doi.org/10.63328/IJCSER-V2I2P3

users who speak diverse language groups. The system achieves a 99% high accuracy rate for sign recognition through the use of a Convolutional Neural Network (CNN) trained on American Sign Language (ASL) dataset. Deep learning implementations in this project boost real-time translation precision while it simultaneously improves gesture identification outcomes. This solution produces dynamic signs through animate animation instead of static sign representations which conventional assistive tools use to provide a user-friendly interactive interface. The database infrastructure maintains process efficiency jointly with user interface components that make the system easy to navigate.

The research tracks down accessibility problems which impact live communication understanding for people who need help with real-time communication. The system achieves smooth communication through its integration of speech recognition and deep learning features alongside real- time animation technology which removes the need for human interpreters. The deployment of machine learning- based sign recognition enables growth within assistive technology while creating possibilities for video communication solutions that use gestural user input.

The web application functions as an intelligent assistive tool which enables communication between members of deaf and mute communities and those who do not use sign language. The system utilizes deep learning together with NLP and computer vision to present a scalable technologically advanced solution which promotes inclusive communication in real-world applications.

2. Related Works

Deep Learning in Sign Language Recognition: A Hybrid Approach

This research develops a deep learning algorithm which recognizes word gestures in order to enhance live sign language detection. The authors describe the obstacles in building independent continuous sign Modeling systems which need to handle the differences in signing speed and duration[1]. Two different deep learning-based techniques comprise the proposed hybrid system to boost sign language recognition precision over continuous periods.

AI-Based Real-Time Speech-to-Text to Sign Language Translation

The review explores AI methods for real-time speech-to-text sign language translation solutions which address communication barriers for hearing impaired patients under COVID-19 circumstances. Researchers have noted sparse evidence about AI and machine learning implementation in this field especially in Africa while suggesting to develop an AI real-time translation tool for South African languages [2]. Recent Advances on Deep Learning for Sign Language Recognition. This research paper evaluates deep learning- based sign language recognition by examining the recent advancements with their challenges and identified opportunities throughout the last five years. The article discusses multiple important points about sign data acquisition technologies and datasets and evaluation ways and different neural network types. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) achieve promising results in fingerspelling and isolated sign recognition according to the study but the research also addresses continuous sign language recognition challenges [3]. A modified deep learning network performs recognition of sign language signals. Researchers use convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to perform deep learning functions that detect associated indicators of signs. This investigation strives to enhance sign language recognition capabilities through deep learning network modifications which strengthen appraisal of sign gestures [4].

A Deep Learning Web Application for Sign Language Recognition.The project concentrates on creating a deep learning model which detects isolated sign language signals before implementing them through a web application deployment. A group of scientists developed an open platform which applies deep learning analytics to perceive sign language movements in real-time for improved hearing- impaired communication [5].

SignNet: A Deep Learning Architecture for Accurate Sign Language Recognition from Images. Through its deep learning design SignNet provides a mechanism that detects sign language within image content precisely. The model draws attention to hand gesture spatial features which enables higher accuracy in recognition while presenting a good option for image-based sign language interpretation [6]. A deep learning system determines sign language communications. An assessment of deep learning approaches for word recognition and classification functions in different sign language video frames forms the focus of this research. A research analysis investigates different deep learning system models to identify their processing efficiency when recognizing sign language gestures recorded in video sequences [7]. A review of research has been performed which focuses on the integration of machine learning with image processing and artificial intelligence to achieve sign language recognition and interpretation. The analysis presents methods to combine these technologies for enhancing system accuracy and efficiency in sign language interpretation and demonstrates how artificial intelligence advances these possibilities in the field [8].

Vol. 2, Issue 2, 2025, https://doi.org/10.63328/IJCSER-V2I2P3

ASL Champ!: A Virtual Reality Game with Deep-Learning Driven Sign Recognition. ASL Champ! provides a virtual reality gaming experience which enhances American Sign Language learning through virtual real-time feedback for students. Using deep learning models for sign recognition the game presents users an interactive platform which lets them practice ASL skills through simulation [9]. A Machine Learning-Driven Web Application for Sign Language Learning. The presented research develops a web platform which implements machine learning to facilitate sign language education. A learning tool for sign language students will be developed by creating algorithms which properly detect expressed signs and transform them into text or spoken language [10]. Artificial Intelligence Technologies for Sign Language The authors reviewed state-of-the-art sign language techniques for capturing, recognition, translation and representation along with their strengths and weaknesses [11]. American Sign Language Recognition and Conversion The authors of 2023 developed hand gesture recognition software by implementing a CNN within an advanced neural network model to enhance ASL gesture recognition [12].

Machine Learning-Driven Web Application for Sign Language Learning. The authors applied their previous sign language recognition methods to web deployment thus enhancing accessibility and usability of communication through sign language (2024) [13]. Users can implement Signtalk as a sign language translation tool to convert signed gestures into spoken text through its neural network mechanism. A system created by detects hand gestures of speech-impaired people to generate speech alongside text representations for better communication access [14]. Indian Sign Language Recognition Using Mediapipe Holistic developed a robust system to convert Indian Sign Language to text or speech, comparing CNN and LSTM models for recognizing static and gesture sign languages [15].

Dataset

A system that performs Sign Language interpretation relies on Artificial Intelligence in combination with Machine Learning resources. The training of Convolutional Neural Network (CNN) models for this project utilizes data from the American Sign Language (ASL) dataset which was obtained through Kaggle. The Japanese dataset houses a complete assortment of ASL hand gesture pictures showing all letters together with digital symbols and frequent vocabulary signs.

The data set maintains its focus on deep learning model training that accurately determines and categorizes sign language gestures. The dataset contains images that belong to different categories which represent individual signs and letters in the database. The RGB color format used for the images provides high-quality features during extraction. Image normalization and resizing and augmentation gains are used for model generalization and overfitting prevention. The dataset splits into three sections to enhance classification performance where training data occupies 80% and validation data amounts to 10% and testing data fills the remaining 10%.

Figure.1: Explanation of the MVC Controller for PMS View Layer

The CNN model obtains spatial capability from hand gestures which enables precise visual-sign-to-text mapping. Each sign has enough available examples in a balanced dataset which enables solid learning operations. This dataset enables the system to reach 99% accuracy which establishes it as a dependable tool for real-time sign language translation together with communication support.

3. System Design and Models

Convolutional Neural Networks (CNNs) brought revolutionary changes to deep learning techniques which excel at identifying images and objects and analyzing video content. Neural networks known as CNNs serve a specialized purpose to understand visual spatial structures in image and video data thereby maintaining high performance in frame recognition tasks. This part discusses basic to advanced CNN frameworks together with their components and operational processes which demonstrate their distinctive capability to extract elements from visual information.

Basic Architecture of CNN **:** A standard CNN contains three essential types of structures: convolutional layers, pooling layers as well as fully connected layers. The network utilizes multiple layers that perform feature reduction and dimension decrease prior to making predictions through learned visual patterns.

Convolutional Layers **:** A CNN mainly operates through its central convolutional layer that utilizes filters to process input images. The small matrix filters of a CNN operate as sliding components that search for specific features including edges, corners and textures when scanning through images. A feature map emerges from the convolution operation that detects particular image patterns. During learning the filters develop in training and they come in different dimensions yet 3x3 and 5x5 represent standard filter sizes.

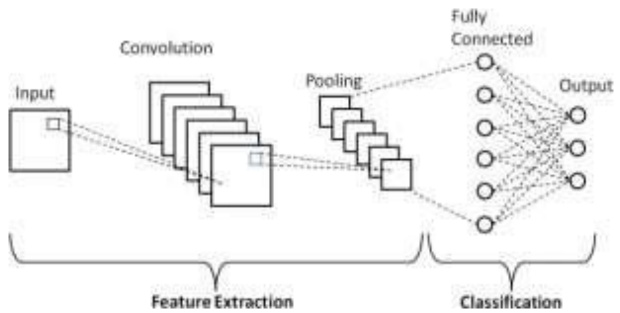

Figure.2 CNN architecture

Activation Function: The model receives an activation function after performing convolution to introduce non-linear characteristics. The ReLU (Rectified Linear Unit) activation function delivers the most efficient performance because it facilitates sparse activation and speeds up learning processes and reduces gradient decay issues.

Pooling Layers: Pooling layers decrease spatial feature map dimensions and reduce parameter counts as well as computational requirements. Both Max pooling along with Average pooling stand among the most frequently employed types of pooling methods. During max pooling the network selects the maximum value from a set of values but average pooling calculates the average of those values. Model overfitting reduction occurs through this layer because it introduces spatial invariance that allows models to focus on key characteristics.

Fully Connected (FC) Layers: A one-dimensional vector is generated after applying multiple convolutional and pooling layers to the input before fully connected layers assess it. The weights of the final classification or regression tasks are assigned by these layers after extracting features from previous steps. Classification models use the softmax activation function in their output layers because it transforms the resultant values to probabilities for every distinct class.

4. Proposed Methodology

User Registration and Login: Users can reach the platform because it features a registration and login system. The built user authentication module utilizes Django platform.

Registration: All users must set up an account by adding their essential details including their email and username and their chosen password. After the registration process users will receive validation instructions through email to begin using their new account.

Login: Users who finish activating their account can access the platform through their username and password combination. When users forget their login details they can activate password reset from the forgot password link.

Access to Dashboard: User authentication leads to a dashboard redirect where users find access to platform main features after successful login occurs.

User Dashboard: All available features of the platform can be found through the user dashboard which provides direct access to them. Users who log in will encounter three principal options including text signing conversion and speech transcription together with translation features.

The dashboard supports an intuitive interface so users can promptly access the Text-to-Sign Language and Speech-to-Text modules through a user-friendly system design.

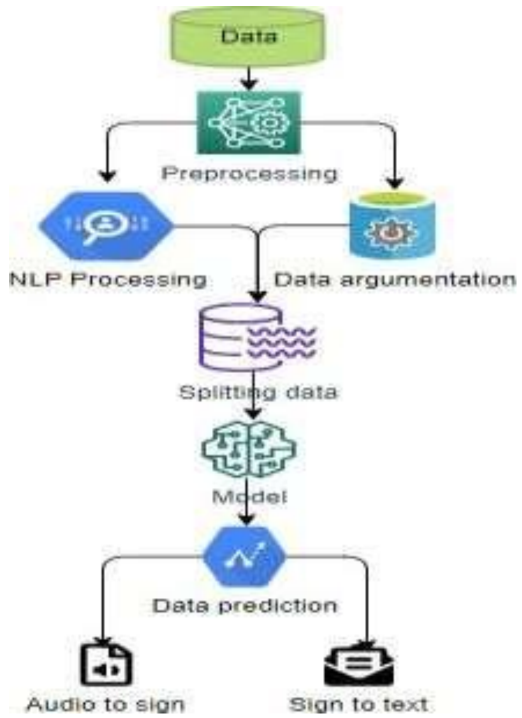

Figure 3 Proposed Work Flow

Text-to-Sign Language Conversion Module

Through this platform users can submit text materials which then get processed into animated sign language gesturing through the system.

Text Input: The program has available text fields whereby users can input their text. After the text entry the system converts it through MediaPipe into animated sign

Vol. 2, Issue 2, 2025, https://doi.org/10.63328/IJCSER-V2I2P3

language gestures for visual display.

Animation Display: Users will benefit from seeing animations of sign language on the display which demonstrate the visual representation of their input text.

Speech-to-Text Module: Users can obtain text transcription from their speech through the integrated speech-to-text module by using a microphone.

Voice Input: A microphone icon activates the voice input when users click it. Through implementation of the Speech Recognition library users can achieve text transcription from speech which happens instantly in real-time.

Text Output: Users can easily observe their spoken words because the transcription displays immediately on the screen.

Text Translation into Indian Languages

The text input method provides access to language translation for various Indian languages after the text becomes available through typing or speech-to-text transcription.

Language Selection: Users of the system have the ability to select any of multiple translation languages including Hindi, Tamil and Bengali and more.

Google Translator API: An API tool from Google Translator enables the system to translate content. The screen display will present the translated text to enable communication between people who speak different languages.

The web application offers a simple, intuitive interface for facilitating communication between deaf, mute, and sign language users. It integrates multiple technologies like CNN, Media Pipe, Speech Recognition, and Google Translator API to provide real-time sign language conversion, speech- to-text functionality, and language translation. By focusing on accessibility and ease of use, the platform ensures that individuals can communicate effectively, regardless of their hearing or speaking ability.

6. Results and Evaluation

The current project provides a good basis for filling the communication gap between deaf, mute, and sign language users, but there are many opportunities for improvement and expansion. One key area for improvement is the integration of real-time video-based sign language recognition.

This will allow the system to interpret directly sign language from live video feeds. This may offer a natural and dynamic form of communication that can be well fitted in face-to-face contact. Moreover, it can develop a mobile application to make access of users wider, allowing the service to access with cell phone on the go, with offline functionality especially for places which do not have good internet access.

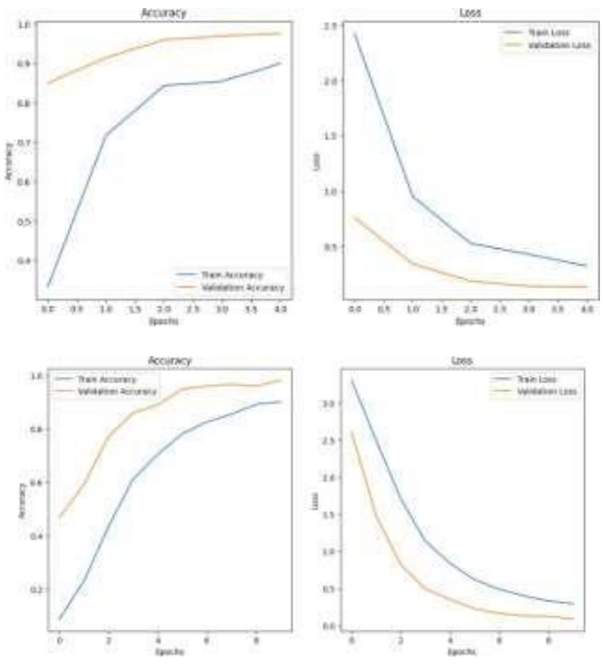

Figure. 4 CNN Accuracy and Loss

Another important future addition would be extending multi- language support. Currently, the system does support a couple of Indian languages, but incorporating more global languages and regional dialects will certainly make the service more inclusive in its appeal for a broader populace. Incorporation of local variants of sign languages from various regional areas would even further improve user inclusivity by catering to differing needs. Furthermore, integrating artificial intelligence techniques that allow the system to continuously learn and improve the performance is another promising area. The system would then be able to adapt and refine its accuracy based on users' interactions in real-time, thus enhancing the all-around experience of a user.

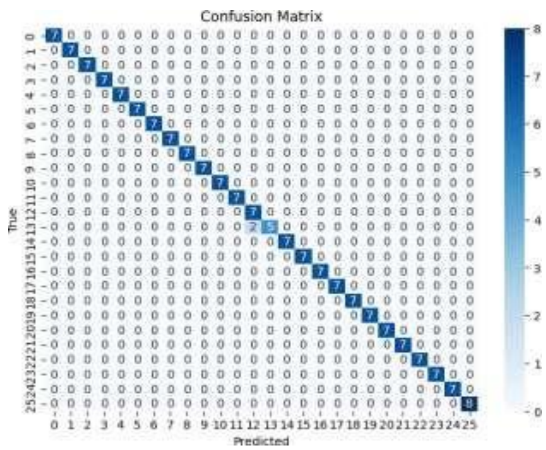

Figure. 5 Confusion Matrix for CNN

Vol. 2, Issue 2, 2025, https://doi.org/10.63328/IJCSER-V2I2P3

The final addition would be voice interaction for sign language users, where they can communicate using gestures while receiving spoken feedback from the system. This would make the experience more interactive and immersive. Exploring these potential advancements will help the platform evolve into a more powerful and scalable solution, paving the way for greater inclusivity in communication for the deaf, mute, and sign language communities globally.

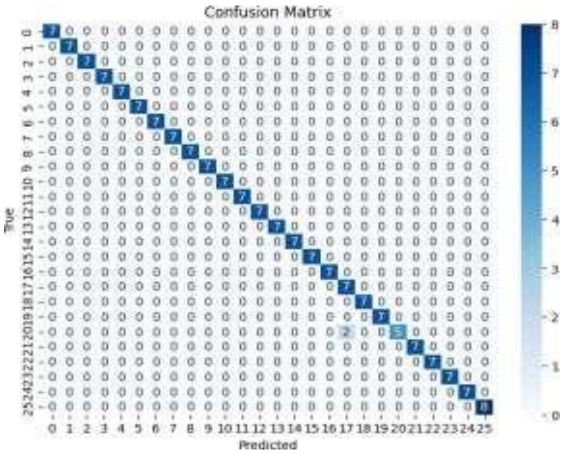

Figure.6 Confusion Matrix for MobileNet

Table.1 & 2 : Measurement of Work flow

Class	Precision & Support		Recall	F1-Score
0	1.00	1.00	1.00	7
1	1.00	1.00	1.00	7
2	1.00	1.00	1.00	7
3	1.00	1.00	1.00	7
4	1.00	1.00	1.00	7
5	1.00	1.00	1.00	7
6	1.00	1.00	1.00	7
7	1.00	1.00	1.00	7
8	1.00	1.00	1.00	7
9	1.00	1.00	1.00	7
10	1.00	1.00	1.00	7
11	1.00	1.00	1.00	7
12	1.00	1.00	1.00	7
13	1.00	1.00	1.00	7
14	1.00	1.00	1.00	7
15	1.00	1.00	1.00	7
16	1.00	1.00	1.00	7
17	0.78	1.00	0.88	7
18	1.00	1.00	1.00	7
19	1.00	1.00	1.00	7
20	1.00	0.71	0.83	7
21	1.00	1.00	1.00	7
22	1.00	1.00	1.00	7
23	1.00	1.00	1.00	7
24	1.00	1.00	1.00	7
25	**1.00**	**1.00**	**1.00**	**8**

Class	Precision & Support		Recall	F1-Score
0	1.00	1.00	1.00	7
1	1.00	1.00	1.00	7
2	1.00	1.00	1.00	7
3	1.00	1.00	1.00	7
4	1.00	1.00	1.00	7
5	1.00	1.00	1.00	7
6	1.00	1.00	1.00	7
7	1.00	1.00	1.00	7
8	1.00	1.00	1.00	7
9	1.00	1.00	1.00	7
10	1.00	1.00	1.00	7
11	1.00	1.00	1.00	7
12	1.00	1.00	0.88	7
13	1.00	0.71	0.83	7
14	1.00	1.00	1.00	7
15	1.00	1.00	1.00	7
16	1.00	1.00	1.00	7
17	0.78	1.00	1.00	7
18	1.00	1.00	1.00	7
19	1.00	1.00	1.00	7
20	1.00	1.00	1.00	7
21	1.00	1.00	1.00	7
22	1.00	1.00	1.00	7
23	1.00	1.00	1.00	7
24	1.00	1.00	1.00	7
25	**1.00**	**1.00**	**1.00**	**8**

Table. 3: Comparison Table

Model	Loss	Accuracy	Val loss	Val Accuracy
CNN	0.2951	0.8911	0.0843	0.9017
Mobilenet	0.3222	0.9096	0.1351	0.9755

8. Conclusion and Future Scope

This project demonstrates an innovative web solution that functions to link communication channels for sign language and speech deaf and mute individuals thus creating more accessible spaces for all users. The system delivers real-time translation solutions through its integration of CNNs and MediaPipe and Speech Recognition and Google Translator API.Sign language recognition using CNN-based models reaches exceptional accuracy levels and hand tracking in MediaPipe produces natural animations that improve comprehension for viewers. The system provides convenient communication through its speech-to-text functionality that turns vocal expressions into written text before it translates them into various Indian languages to tackle language Users will find the system easy to use because developers implemented a user-friendly design including customizable options and animation speed controls and multilingual access. The proposed system demonstrates promising capability to advance communication capabilities among deaf and mute users together with people who do not understand sign language. Through this system users can maintain dialogue without requiring human help thus creating more opportunities for social integration along with user empowerment. Through ongoing development the web application serves as an inclusive tool to make communication more accessible and shows potential to drive more assistive technology advances for deaf and mute populations.

References

[1]. A. M. Buttar, U. Ahmad, A. H. Gumaei, A. Assiri, M. A. Akbar, and B. F. Alkhamees, "Deep Learning in Sign Language Recognition: A Hybrid Approach for the Recognition of Static and Dynamic Signs," Mathematics 2023, Vol. 11, Page 3729, vol. 11, no.17, p. 3729, Aug. 2023, doi: 10.3390/MATH11173729.

[2]. M. Madahana, K. Khoza-Shangase, N. Moroe, D. Mayombo, O. Nyandoro, and J. Ekoru, "A proposed artificial intelligence-based real-time speech-to-text to sign language translator for South African official languages for the COVID-19 era and beyond: In pursuit of solutions for the hearing impaired," The South African Journal of Communication Disorders, vol. 69, no. 2, p. 915, 2022, doi: 10.4102/SAJCD.V69I2.915.

[3]. Y. Zhang and X. Jiang, "Recent Advances on Deep Learning for Sign Language Recognition," Computer Modeling in Engineering & Sciences, vol. 139, no. 3, pp. 2399–2450, Mar. 2024, doi: 10.32604/CMES.2023.045731.

[4]. A. Baihan, A. I. Alutaibi, M. Alshehri, and S. K. Sharma, "Sign language recognition using modified deep learning network and hybrid optimization: a hybrid optimizer (HO) based optimized CNNSa-LSTM approach," Scientific Reports 2024 14:1, vol. 14, no. 1, pp. 1–22, Oct. 2024, doi: 10.1038/s41598-024-76174-7.

[5]. A. Sharma, D. Guo, A. Parmar, J. Ge, and H. Li, "Promoting Sign Language Awareness: A Deep Learning Web Application for Sign Language Recognition," ACM International Conference Proceeding Series, pp. 22–28, Nov. 2024, doi: 10.1145/3695719.3695723/ASSET/8EA1D820-783D-43F3-9B01 B63D8A32007A/ASSETS/GRAPHIC/ICDLT2024- 4-FIG6.JPG.

[6]. K. Sahu, T. Saraswat, and R. Sawhney, "SignNet: A Deep Learning Architecture for Accurate Sign Language Recognition from Images," 2023 14th International Conference on Computing Communication and Networking Technologies,ICCCNT 2023, 2023, doi: 10.1109/ICCCNT56998.2023.10306591.

[7]. M. Mahyoub, F. Natalia, S. Sudirman, and J. Mustafina, "Sign Language Recognition using Deep Learning," Proceedings - International Conference on Developments in eSystems Engineering, DeSE, vol. 2023-January, pp. 184–189, 2023, doi: 10.1109/DESE58274.2023.10100055.

[8]. F. M. Najib, "Sign language interpretation using machine learning and artificial intelligence," Neural Comput Appl, vol. 37, no. 2, pp. 841–857, Nov. 2024, doi: 10.1007/S00521-024-10395-9/TABLES/6.

[9]. M. S. Alam et al., "ASL Champ!: A Virtual Reality Game with Deep-Learning Driven Sign Recognition," Dec. 2023, Accessed: Feb. 04, 2025. [Online]. Available: https://arxiv.org/abs/2401.00289v1

[10]. H. Orovwode, O. Ibukun, and J. A. Abubakar, "A machine learning-driven web application for sign language learning," Front Artif Intell, vol. 7, p. 1297347, Jun. 2024, doi: 10.3389/FRAI.2024.1297347/BIBTEX.

[11]. I.Papastratis, C. Chatzikonstantinou, D. Konstantinidis, K. Dimitropoulos, and P. Daras, "Artificial Intelligence Technologies for Sign Language," Sensors 2021, Vol. 21, Page 5843, vol.21, no. 17, p. 5843, Aug. 2021, doi:10.3390/S21175843.

[12]. [A. Bailur, Y. Limbachia, M. Shah, H. Shah, and Prof. A. Kachare, "American Sign Language Recognition and its Conversion from Text to Speech," Int J Res Appl Sci Eng Technol, vol. 11, no. 9, pp. 1533–1538, Sep. 2023, doi: 10.22214/IJRASET.2023.55871.

[13]. H. Orovwode, O. Ibukun, and J. A. Abubakar, "A machine learning-driven web application for sign language learning," Front Artif Intell, vol. 7, p. 1297347, 2024, doi: 10.3389/FRAI.2024.1297347.

[14]. C. U. Bharathi, G. Ragavi, and K. Karthika, "Signtalk: Sign Language to Text and Speech Conversion," 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation, ICAECA 2021, 2021, doi: 10.1109/ICAECA52838.2021.9675751.

[15]. V. G and K. Goyal, "Indian Sign Language Recognition Using Mediapipe Holistic," Apr. 2023, Accessed: Feb. 04, 2025. [Online]. Available: https://arxiv.org/abs/2304.10256v1