



Empowering Communication:

A Web Application for Deaf, Mute, and Sign Language Interpretation

**Konda Hari Krishna ¹ , B. Suchitha ² , B. Pravalika ³ , E Deepika ⁴ ,
G Jaahnavi ⁵ , Teklu Urgessa ⁶**

¹⁻⁵ Department of Computer Science and Engineering , School of Computing , Mohan Babu University , Tirupati , Andhra Pradesh, India - 517102 , khk396@gmail.com

⁶ Department of Computer Science and Engineering, Adama Science and Technology University, Adama, Ethiopia ; teklurgessa2009@gmail.com

* Corresponding Author : Konda Hari Krishna ; 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 know 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 collected with text- to- Sign language conversion and speech-to-text capabilities towards 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 Speech Recognition 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 model reaches 99% accuracy for sign recognition leading to accurate communication. The application creates a manageable 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, and utilizes Google Translator API.

1. Introduction

The ability to communicate is essential to 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 & 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 approaches to change spoken words into written text. The text processing through Google Translator API enables language conversion to different Indian regional languages thereby broadening access for 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 DL 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 not at all used Sign language. The system utilizes DL together with NLP and computer vision to present a scalable technologically advanced solution which promotes inclusive communication in real-world applications.

2. Related Work

2.1. DL in Sign Language Recognition: A Hybrid Approach

This work develops a DL 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. Two different DL-based techniques comprise the proposed hybrid system to boost Sign language recognition precision over continuous periods [1].

2.2. AI-Based Sign language to speech-to-text Translation

The review discusses AI-based approaches to real-time speech-to-text Sign language translation solutions that can solve the communication barrier in hearing impaired patients in COVID-19 situations. In their study, researchers have observed about the lack of data regarding AI and machine learning application in this area particularly in Africa as they propose to create an AI real-time translation tool related to the South African languages. [2].

2.3. Recent Progress in DL based Sign language Recognition

This research paper analyses the DL-based Sign language recognition by analyzing the current developments with its challenges and opportunities that have been realized in the past five years. Several central issues are raised in the article concerning the technology of sign data acquisition and data sets and modes of evaluation and the various types of neural networks. The study confirms that CNNs(CNNs) and Recurrent Neural Networks (RNNs) are promising in fingerspelling and isolated sign recognition but the research also covers continuous Sign language recognition issues. [3].

2.4. An adapted DL network carries out Sign language signal recognition

DL functionalities based CNNs and recurrent neural networks (RNNs) are used by researchers in order to identify related indicators of signs. This research aims at maximizing Sign language recognition at the expense of DL network adjustments that reinforce the evaluation of sign gestures. [4].

2.5. DL Sign language Recognition Web Application

The project will focus on developing a DL tool that identifies individual signals of Sign language and then realizes them in the form of a web application. One of the groups of scientists created a general platform that uses DL analytics to view the Sign language movements in real time and enhance communication between the hearing-impaired [5].

2.6. SignNet

A DL Architecture of Detailed Sign language Recognition on Image. The SignNet is a system that identifies Sign language in the image material with accuracy because of its DL design. This model gives emphasis

to spatial features of hand gestures that allows greater accuracy in recognition and leaves a good alternative of image-based Sign language decoding [6].

2.7. Sign language communications are determined by a DL System

The scope of the research is the evaluation of DL solutions to word recognition and classification tasks in various Sign language videos frames. A study examines various models of DL system to determine the processing effectiveness of the systems in detecting hand movements in video recordings that represent Sign language [7]. The research analysis has been conducted that revolves around combining machine learning and image processing with artificial intelligence to attain Sign language recognition and interpretation. The analysis includes ways of incorporating these technologies to increase the accuracy and efficiency of the system in interpreting Sign language and shows how AI expands these opportunities in the sphere [8].

ASL Champ! A Video Game of Virtual Reality with Sign Recognition based on Deep-Learning. An example is ASL Champ!, a virtual reality game that supplements student learning in American Sign language because it gives students virtual real-time feedback. Through sign recognition the game is an interactive platform based on DL that allows users to train in ASL skills via simulation. [9]. Machine Learning-Powered Sign language Learning Web Application. The given study develops a web app that uses machine learning to support the learning of the Sign language. Sign language students will have a learning tool created that will involve creation of algorithms that adequately identify expressed signs and encode them into either text or spoken language.

Sign language Artificial Intelligence Adult Sign language systems The authors of the reviewed state-of-the-art Sign language capture, recognition, translation and representation systems together with their advantages and disadvantages [11]. American Sign language Recognition and Conversion the authors of 2023 created the software of recognition of hand gestures through the implementation of a CNN into an advanced neuro model to improve the recognition of gestures in ASL. [12]. Web Application with Sign language Learning with the use of Machine Learning. The authors used the Sign language recognition techniques they developed previously to web deployment thereby increasing the accessibility and usability of communication using Sign language (2024). [13].

Signtalk can be used by users as a Sign language translation system to transform signed language into the spoken word by the use of its neural network system. One such system is known as detects the hand gestures of

speech-impaired individuals in order to produce speech along with text representations to aid easier communication access. [14]. Indian Sign language Recognition Using MediaPipe Holistic created a strong platform in transforming [15] the Indian Sign language to text or speech, comparison of CNN and LSTM models in recognizing both static and gesture languages of the Sign.

2.8. Dataset

CNNs models to be trained on this project are based on the American Sign language (ASL) dataset that was acquired on 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 DL 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 Colour 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.



Figure.1 ASL dataset samples

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%. 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 database can provide the system with 99% accuracy that makes it a reliable system of real-time Sign language translation along with communication assistance.

3. System Design and Models

CNNs brought revolutionary changes to DL 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.

3.1. Basic Architecture of CNN

A basic CNN consists of three key features of structures: convolutional layers, pooling layers and fully connected layers. The network employs the use of several layers that eliminate features and decrease of dimensions before making predictions based on 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.

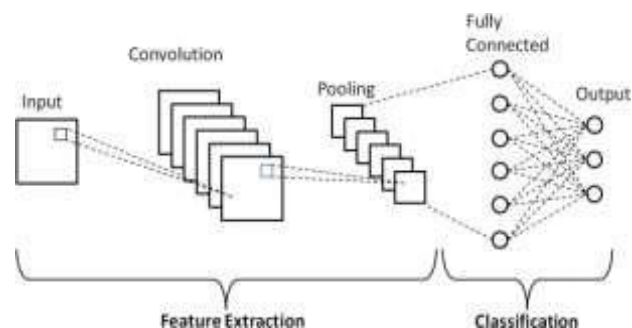


Figure.2 CNN Architecture

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. In the process of max pooling the network chooses the highest value out of a pool of values whereas in average pooling the average value is calculated. 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

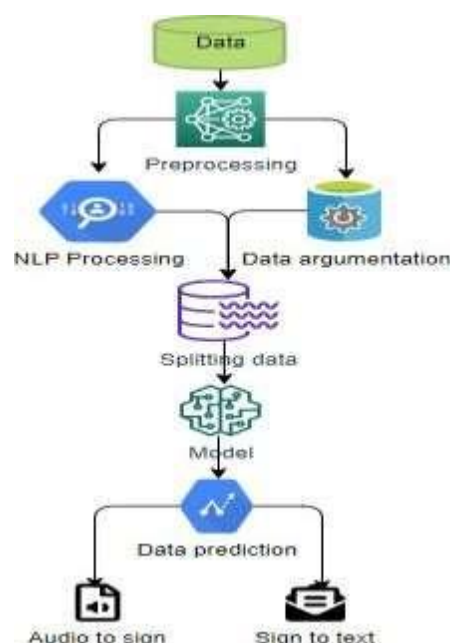


Figure. 3 Proposed Work Flow

4.1. 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.

4.2. 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. Text-to-Sign language Conversion Module Through this platform user 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 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.

4.3. 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, MediaPipe, 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.

5. Results and Conclusion.

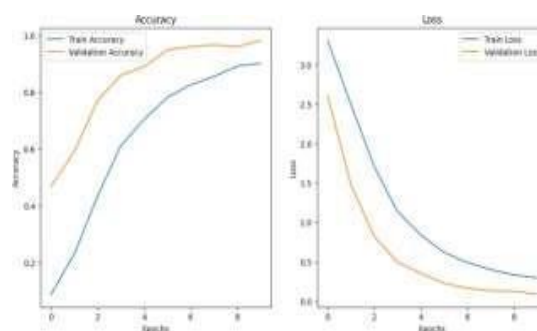


Figure. 4 CNN Accuracy and Loss

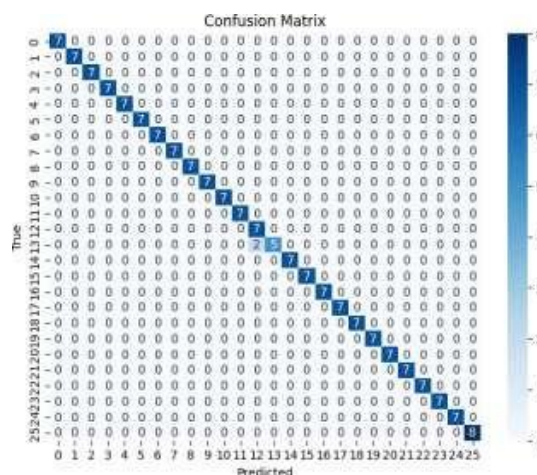


Figure. 5 Confusion Matrix for CNN

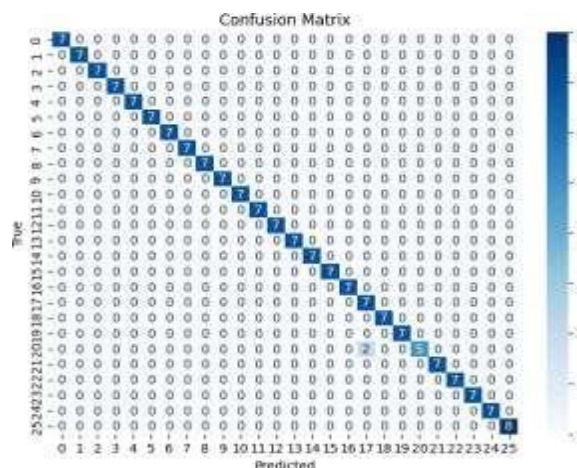


Figure. 6 Confusion Matrix for MobileNet

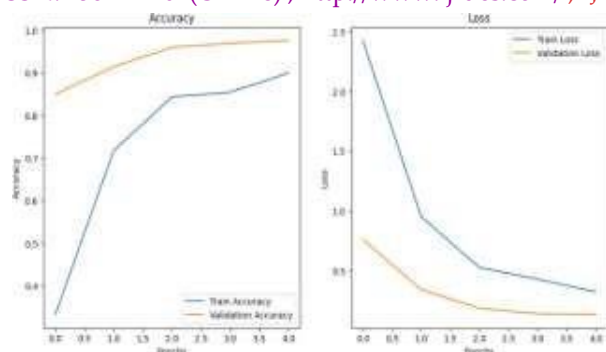


Figure. 7 Training and validation loss MobileNet

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

Figure. 8 MobileNet Model

Table. 1 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

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

Figure. 9 CNN Model

7. 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.

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.

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. 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.

Reference

- [1]. A. M. Buttar, U. Ahmad, A. H. Gumaie, 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, "Signlanguage recognition using modified deep learning network and hybrid optimization: a hybrid optimizer (HO) based optimized CNNs- 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 Signlanguage Awareness: A Deep Learning Web Application for Signlanguage 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. Bailer, 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>

Author Contribution

KHK, BS, BP : Conceptualization, drafting, data collection and analysis, Methodology, and Supervision;

ED, GJ : Methodology, Writing and Data Collection, and Editing

Declaration

Conflicts of Interest: The authors declare no conflict of interest.

Author Contribution: All authors wrote the main manuscript text and also consent to the submission.

Plagiarism : Similarity Check - 8 % , AI Plagiarism : * %

Ethical approval: Not applicable.

Consent to Participate: All authors consent to participate.

Funding: Not applicable, and No funding was received

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Personal Statement: We declare with our best of knowledge that this research work is purely Original Work and No third party material used in this article drafting. If any such kind material found in further online publication, we are responsible only for any judicial and copyright issues.

Acknowledgements: We thank everyone who inspired our work.