iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** III **Month of publication:** March 2025

DOI: https://doi.org/10.22214/ijraset.2025.67381

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue III Mar 2025- Available at www.ijraset.com

Real Time Detection and Conversion of Gestures to Text and Speech to Sign System

Mrs. D. Srivalli[1], Komma Mounika[2], Odugu Hemanth[3], Gompa Lokesh[4], Merugu Ratnakumar[5]
[1]Assistant Professor, [2,3,4,5]Scholars, Department of Computer Science & Engineering, Dhanekula Institute of Engineering and Technology, A.P, India

Abstract: Millions of individuals throughout the world use sign language as their primary form of communication, particularly those who are deaf or hard of hearing. But because not everyone is proficient in sign language, it might be difficult to communicate and restrict access to information for sign language users This project represents a groundbreaking integration of technology in the field of communication and accessibility One of the many challenges is the ability to communicate with others through sign language. The goal of the system is to convert speech to text using NLP(natural language processing),Mediapipe for hand landmark detection and CNN for sign classification

Keywords: sign language recognition, NLP, CNN, Mediapipe, speech to sign, Sign to text.

I. INTRODUCTION

Sign language is used by deaf and mute people to to communicate with others. It is a form of visual communication. It is a fully developed language that uses its own rules of grammar and dictionaries. It uses hand gestures, body movements and even facial expressions and speeches to convey thoughts. In this project we are using hand gestures.It does not use traditional word patterns. With the rising number of people suffering from hearing loss, it is crucial to find ways to bridge the communication gap between the hearing and non-hearing population. To address this issue, we present a new system for converting Sign Language into text format using computer vision and machine learning techniques.. This system focuses on Indian Sign Language and utilizes the Media Pipe Holistic Key points for hand gesture recognition.Hence it is a real time AI based translation system.

Converting speech into sign language using NLP and a gesture mapping system.Translating sign language into text using Mediapipe,CNN based gesture recognition

II. LITERATURE REVIEW

[1] "Integrating NLP and Computer Vision for Effective Sign Language Translation"**:** This paper explores the intersection of Natural Language Processing (NLP) and Computer Vision in the development of sign language translation tools. We present a comprehensive review of current methodologies, focusing on the integration of NLP for understanding contextual nuances and computer vision for accurate gesture recognition. The study highlights the challenges and advancements in real-time translation systems, providing insights into future research directions for enhanced communication tools for the deaf and hard-of-hearing community.

[2]"Advancements in Real-Time Speech-to-Sign Language Systems": In this survey, we examine recent technological advancements in converting real-time speech to sign language. The paper reviews various machine learning models and their effectiveness in interpreting spoken language nuances and translating them into accurate sign language gestures. By analyzing case studies and experimental results, the paper sheds light on the progress in assistive communication technologies, discussing the implications for accessibility and inclusivity in diverse societal contexts.

[3]Real time hand tracking using mediapipe”:: In this survey, we examine that mediapipe can detect 21 hand landmarks per frame with minimal computational overhead,making it idea for real time applications.

[4] "A Comparative Analysis of CNN based Gesture Recognition": Our survey paper presents a detailed comparative analysis of gesture recognition technologies used in sign language translation. It covers a range of techniques from traditional image processing to advanced deep learning approaches. The paper evaluates the accuracy, speed, and reliability of these technologies, providing a critical assessment of their application in real-world scenarios for the deaf and hard-of-hearing.

III. PROPOSED SYSTEM

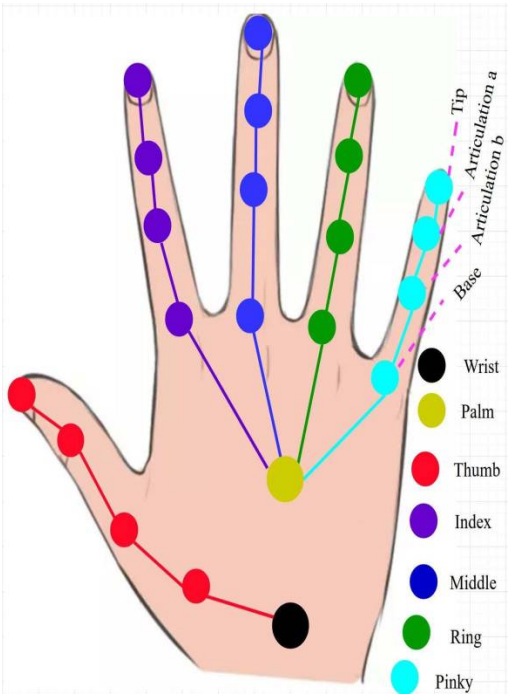
The system is designed to translate between spoken language and sign language in real time.it consists of two main modules:

A. Speech to Sign Conversion

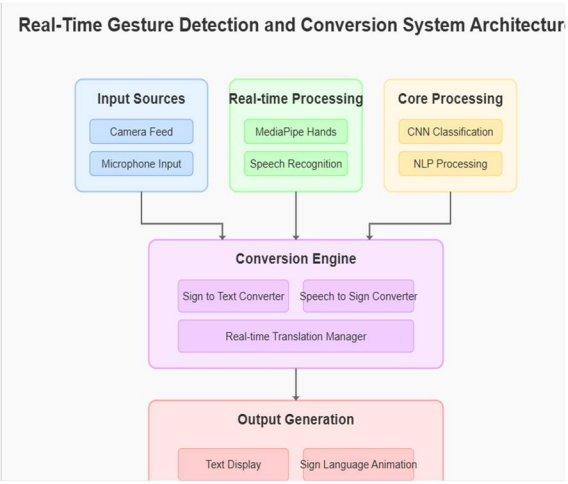
Converts spoken words into corresponding sign gesture.Using microphone the user gives the audio.An NLP based models transcribes the audio into text.The transcribed text is cleaned.The text is restructured to reflect sign language grammar,which often differs from spoken language syntax.Each word is mapped to its corresponding sign gesture.

B. Gesture to Text Conversion

Here the process is to converts sign gesture to text.we are using mediapipe for hand tracking.which detects and tracks 21 landmarks in every video.These landmarks represents the positions and movements of the hands and fingers.CNN is used for gesture classification.The normalized landmark coordinates are fed into a CNN.Recognized sign are sequenced together to form words and sentences.



Architecture



IV. METHODOLOGY

A. Speech to Sign Language Conversion

- Step1: speech recognition Convert spoken words into text using asr. Preprocess the text by removing noise.
- Step2:text processing&tokenization Tokenize the text into individual words. Apply NLP techniques to identify key components.simplify complex sentences into a form that aligns with sign language grammar.
- Step3: NLP and Database matching: Map words to sign language.
- Step4: shows gesture animation

B. Sign to Text

- Step1:Hand gesture and pose detection Capture real time video using camera.
- Step2:use CNN for gesture classification Convert detected hand gesture into a feature vector using CNN. Train the model on a sign language dataset to classify each sign.
- Step 3:Text generation Convert sign to text. And display translated text on screen.

V. RESULT

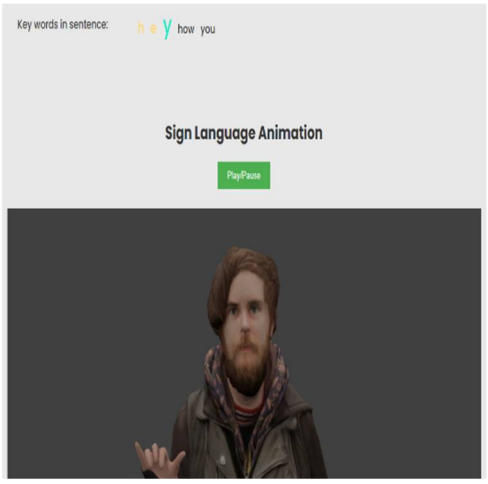


Fig.1

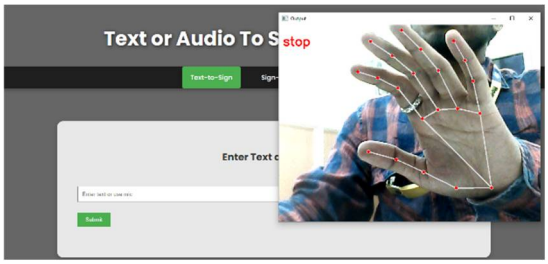


Fig.2

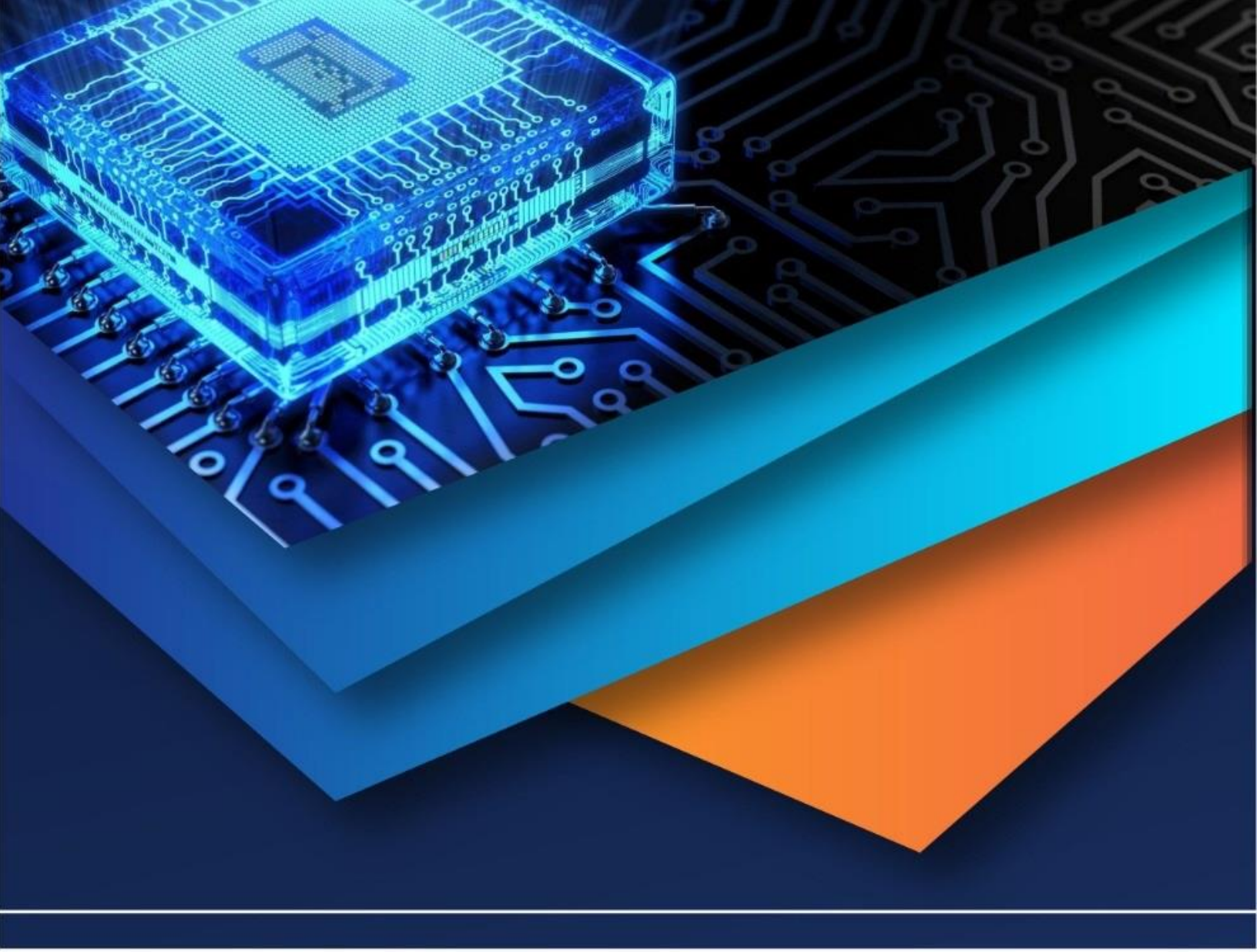
VI. CONCLUSION

This project has shown promising results,in converting spoken language to sign and sign to text.By leveraging NLP techniques such as POS tagging,lemmati-zation etc ,the system can preprocess the input audio and extract the meaningful features.The resulting text is then mapped to sign language gesture which can viewed by deaf people.

This project successfully integrates several complex technologies, including natural language processing (NLP), speech-to-text conversion, gesture recognition using Media Pipe, movements of the hands and fingers.CNN is used for gesture classification and dynamic sign language animation. The result is a versatile and user-friendly platform that facilitates real-time communication between individuals who use sign language and those who do not, bridging a critical gap in interpersonal interactions.

REFERENCES

[1] Koller, O., Ney, H., & Bowden, R. (2016). "Deep Learning of Mouth Shapes for Sign Language." In Proceedings of the IEEE International Conference on Computer Vision Workshops.
[2] Bragg, D., et al. (2019). "Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective." The 21st International ACM SIGACCESS Conference on Computers and Accessibility.
[3] McNeil, M. R. (Ed.). (2016). "Clinical Management of Sensorimotor Speech Disorders." Thieme.
[4] A. S. M. Miah, J. Shin, M. Al M. Hasan, M. A. Rahim, andY. Okuyama, ‘‘Rotation, translation and scale invariant sign wordrecognition using deep learning,’’*Comput. Syst. Sci. Eng*., vol. 44, no. 3,pp. 2521–2536, 2023.
[5] A. S. M. Miah, J. Shin, M. A.M. Hasan, and M.A.Rahim, BenSignNet:Bengali sign language alphabet recognition using concatenated segmenta-tion and convolutional neural network,’’*Appl. Sci*., vol. 12, no. 8, p. 3933,Apr. 2022.
[6] S. S. Rautaray and A. Agrawal, ‘‘Vision based hand gesture recognition forhuman computer interaction: A survey,’’*Artif. Intell. Rev*., vol. 43, no. 1,pp.1–54, Jan. 2015.
[7] Parton, B. S. "Sign Language Recognition and Translation: A Multidisciplined Approach from the Field of Artificial Intelligence." Journal of Deaf Studies and Deaf Education.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY