S.I.G.N. - Sign Interpretation and Gesture Navigation

Sania Kalpesh Pimple
Computer Engineering Department
Atharva College of Engineering,
Mumbai University, India
pimplesania-cmpn@atharvacoe.ac.in

Devashree Rajesh Karande
Computer Engineering Department
Atharva College of Engineering,
Mumbai University, India
karandedevashree-cmpn@atharvacoe.ac.in

Mann Chirag Shukla
Computer Engineering Department
Atharva College of Engineering,
Mumbai University, India
shuklamann-cmpn@atharvacoe.ac.in

Chaitali Parab
Computer Engineering Department
Atharva College of Engineering,
Mumbai University, India
chaitaliparab-cmpn@atharvacoe.ac.in

Prof. Nileema Pathak
Computer Engineering Department
Atharva College of Engineering,
Mumbai University, India
nileemapathak@atharvacoe.ac.in

Abstract - ***The S.I.G.N (Sign Interpretation using Gesture Navigation) project addresses the communication gap between sign language users and non-users, especially in sectors like education, healthcare, and public services.The project places a strong emphasis on Indian Sign Language (ISL) to cater to the communication needs of the hearing-impaired community in India. By tailoring recognition models and sign databases to ISL, the system ensures higher relevance, usability, and cultural alignment for Indian users. It utilizes machine learning, computer vision, and natural language processing (NLP) to translate sign language gestures into meaningful text or speech in real time. The system comprises two key components: (1) Real-time Sign to Sentence Generation, which converts live gestures into accurate, context-aware sentences, and (2) Word to Sign Generation, which displays typed words as corresponding sign gestures for learning and interaction.***

Convolutional Neural Networks (CNNs) handle gesture recognition, while MediaPipe and OpenPose ensure precise pose estimation. NLP techniques enhance grammatical accuracy, and the system achieves over 95% accuracy, delivering reliable performance. A Streamlit-based interface enables smooth real-time interaction, making S.I.G.N a scalable, cost-effective, and inclusive alternative to human interpreters.

Keywords: ***Computer Vision, Deep Learning, Convolutional Neural Networks (CNNs), Machine Learning, Natural Language Processing (NLP), Image Processing, Feature Extraction, Hand Tracking.***

I. INTRODUCTION

Indian Sign Language (ISL) is the primary mode of communication for a significant portion of the deaf and hard-of-hearing community in India. According to the Indian Census and various disability reports, **over 8 million people** in India are deaf or speech-impaired, many of whom rely on ISL for daily communication. Despite its widespread use, ISL lacks universal standardization and is not widely understood by the general population, creating communication barriers in education, healthcare, and employment. The growing demand for interpreters and inclusive technologies highlights the urgent need for ISL-focused solutions like S.I.G.N.

Sign language is a vital mode of communication for individuals with hearing or speech impairments. However, limited public understanding and a shortage of qualified interpreters create significant barriers in areas like healthcare, education, and employment. To address this, we propose a **Real-time Sign Language Interpreter and Gesture Navigation System**, designed to bridge communication gaps and promote inclusivity.

Our system consists of two primary modules: (1) **Real-time Sign to Sentence Generation**, which recognizes live sign language gestures and converts them into grammatically correct sentences using Natural Language Processing (NLP), and (2) **Word to Sign Generation**, which visually translates typed words into corresponding sign language gestures for learning and interaction.

Developed with **Streamlit** for the interface, the system integrates **MediaPipe** and **OpenCV** for gesture and pose recognition, **CNNs** for model training, and **NLP** techniques for accurate sentence formation. Additional libraries include **NumPy**, **Pandas**, **Matplotlib**, **PIL**, **GoogleTranslator (deep_translator)**, and standard utilities such as **random**, **subprocess**, and **warnings**. The solution achieves an accuracy exceeding **95%**, offering a scalable, cost-effective alternative to human interpreters and contributing significantly to accessible communication for the hearing-impaired community.

II. LITERATURE REVIEW

1. Aakash Deep; Aashutosh Litoriya; Akshay Ingole; Vaibhav Asare; Shubham M Bhole; Shantanu Pathak, "Realtime Sign Language Detection and Recognition": The real-time sign language recognition system was developed for recognizing the gestures of Indian Sign Language (ISL). Generally, these sign languages consist of hand gestures. For recognizing the signs, those regions are identified and tracked using the segmentation feature of OpenCV. Then by using Mediapipe,
it captures the landmarks of the hands and the key points of landmarks are stored in an NumPy array. Then they can train the model on it by using TensorFlow, Keras and LSTM. Previous approaches to sign detection and recognition were done by using the Machine Learning Algorithm by training it on the images but now they are using Deep Learning Models to enhance Realtime sign detection and recognition.

2. "Study and Survey on Gesture Recognition Systems" Information Technology Pune, India: Gestures are vital nonverbal signals that enhance human communication through movements, facial expressions, and body posture. They play a significant role in Human-Computer Interaction (HCI) by providing intuitive and accessible ways for users to engage with technology, particularly benefiting those with physical disabilities. Gestures also facilitate communication across language barriers and are essential in sign language, which employs hand movements, body posture, and facial expressions to convey
meaning for the deaf and hard of hearing.This study focuses on the architecture and operation of gesture recognition systems, exploring methodologies for detecting and interpreting gestures, as well as various applications.

3. "Perspective and Evolution of Gesture Recognition for Sign Language: A Review" Luis Esteban-Hernández Aries Research Center, Universidad Antonio Nebrija, 28015: This paper highlights that gestures are a key form of non-verbal communication, where the body conveys specific messages without spoken words. These gestures involve the movement of various body parts—primarily the hands, arms, and facial expressions. Among these, hand gestures are considered the most expressive and commonly used due to their intuitive and natural nature. Studies, including those by communication researcher Albert Mehrabian, suggest that gestures and visual signals can account for up to 55% of the overall impact in conveying emotions during communication, emphasizing their significant role in enhancing spoken messages.

III. PROPOSED SOLUTION

To address the communication gap faced by individuals who rely on Indian sign language, this project proposes the development of a real-time, vision-based Sign Language Interpreter that converts hand gestures into text or speech using computer vision and rule-based gesture recognition techniques and also text to 3D animated videos. Unlike AI-based solutions, which require extensive training datasets and high computational power, this system will provide a lightweight, efficient, and cost-effective alternative.
Our proposed solution consists of the following components:

1. Real-Time Video Capture

○ A **camera** (webcam or external) continuously captures **live video** of the user's hand movements.

○ The video frames are processed to detect and isolate hand gestures.

2. Hand Detection & Tracking

○ The system will utilize **MediaPipe** Hand Tracking and **OpenCV** to identify key hand landmarks, including fingertips, palm position, and finger joints.

3. Gesture Recognition Using Rule-Based Processing

○ Instead of AI or deep learning models, gestures will be recognized using contour detection, convex hull analysis, and fingertip counting.

○ This method ensures fast processing and low computational requirements, making it suitable for real-time applications.

4. Text and Speech Output Generation

○ Once a gesture is recognized, it will be converted into text output and displayed on the screen.

○ A Text-to-Speech (TTS) module will be integrated to vocalize the recognized words, improving accessibility for those who do not understand sign language.

5. User-Friendly Interface

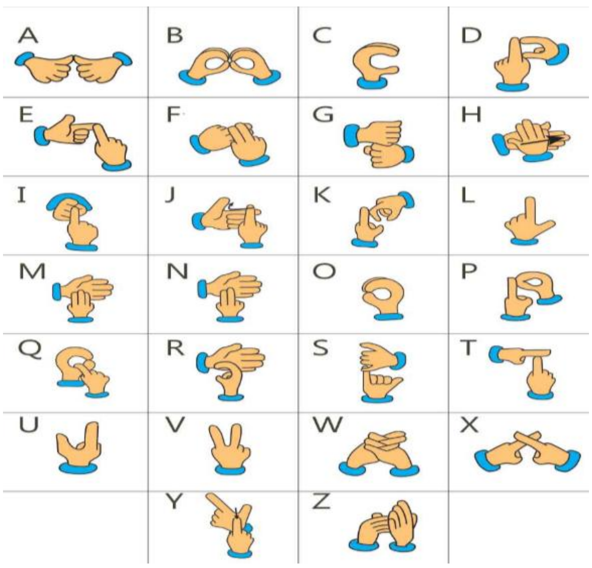
○ The system will feature an interactive and intuitive graphical user interface (GUI) that provides real-time feedback.

○ The interface will allow users to view the translated text, access gesture dictionaries, and interact with the system efficiently.

6. Deployment as a Web-Based

○ The proposed solution will be deployed as a streamlit-based web application, making it accessible via browsers on different available devices.

○ Also reference charts and videos will be available for better understanding.



IV. METHODOLOGY

4.1 Data Collection & Preprocessing:

We use publicly available sign language datasets (e.g., ISL, ASL Dataset). Preprocessing includes resizing images, normalization, and data augmentation (rotation, contrast adjustment) to improve model generalization.

4.2 Hand Gesture Recognition:

For hand gesture recognition, the system employs MediaPipe to detect and track hand key points. These extracted features are then fed into a CNN-based model, which classifies different sign gestures. To ensure an accurate and meaningful translation of sign language, NLP techniques are integrated. Recognized gestures are processed using sequence - to - sequence models and converting them into grammatically correct sentences.

4.3 System Architecture:

The overall system architecture follows a structured pipeline: camera input captures sign gestures, computer vision models process the images, NLP refines the translation, and the final output is displayed as text or converted into speech using a Text-to-Speech (TTS) engine. This ensures a seamless, real-time, and interactive experience for users.

V. IMPLEMENTATION & TECHNOLOGIES USED

5.1 Implementation:

The implementation of the Sign Language Interpreter follows a step-by-step process that integrates various computer vision and software components:

- **Data Acquisition**: The system continuously captures video frames from a connected webcam or external camera device.

- **Preprocessing**: The captured frames are processed using OpenCV to adjust contrast, reduce noise, and improve gesture visibility.

- **Hand Tracking and Feature Extraction**: MediaPipe detects key hand landmarks, and contour analysis is used to determine gesture positions.

- **Gesture Recognition and Mapping**: A predefined rule-based algorithm interprets detected gestures and maps them to sign language symbols.

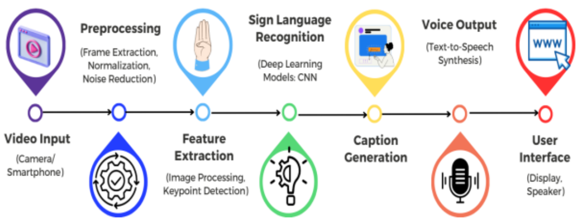


Fig.1

Implementation Flow

- **Output Processing**: The recognized sign is displayed as text on the user interface and optionally converted into speech using TTS.

- **User Interaction**: A simple and intuitive graphical user interface (GUI) allows users to view translations and adjust settings.

5.2 Technologies Used:

- **OpenCV**: It handles video capture, image processing, and contour detection. Essential for preprocessing and extracting hand features from real-time video.

- **MediaPipe**: It provides fast and efficient hand tracking and landmark detection, enabling accurate interpretation of hand gestures through pose estimation.

- **NumPy & Math Libraries**: Used for performing geometric calculations such as distances, angles, and coordinate transformations, which are crucial for accurate gesture classification.

- **Text-to-Speech (TTS)**: It converts recognized text into audible speech.

- **Streamlit**: Powers the project's graphical user interface (GUI), user-friendly real-time communication.

- **Google Translate (deep_translator)**: Enables dynamic translation of recognized text into multiple languages.

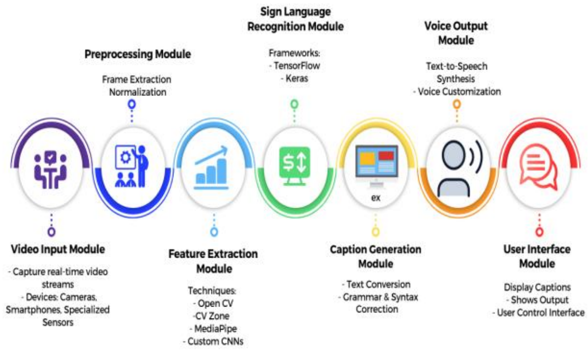


Fig.2: Technical Stack and Libraries

- **Convolutional Neural Networks (CNNs)**: Used for gesture recognition and classification based on visual input, enabling high-accuracy predictions.

VI. DESIGN, WORKING AND PROCESSES

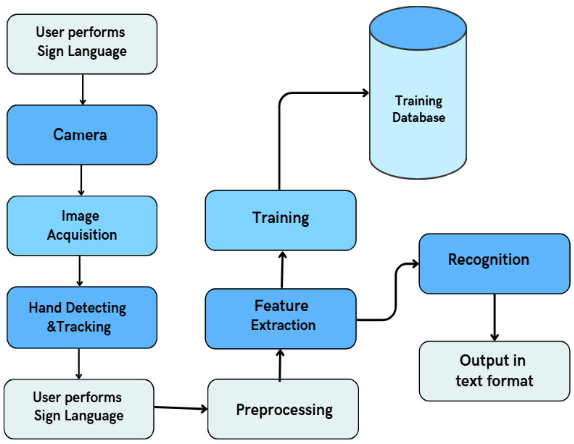


Fig.3: Real-time Sign Detection

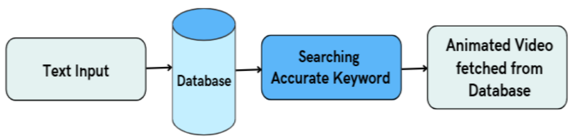


Fig.4: Word to Sign Generation

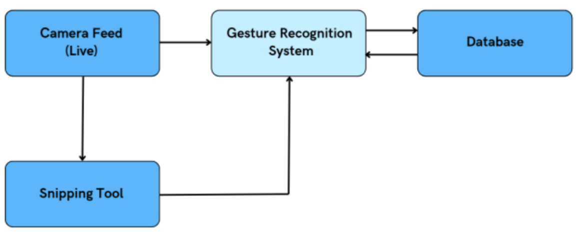


Fig.5: Camera Workflow

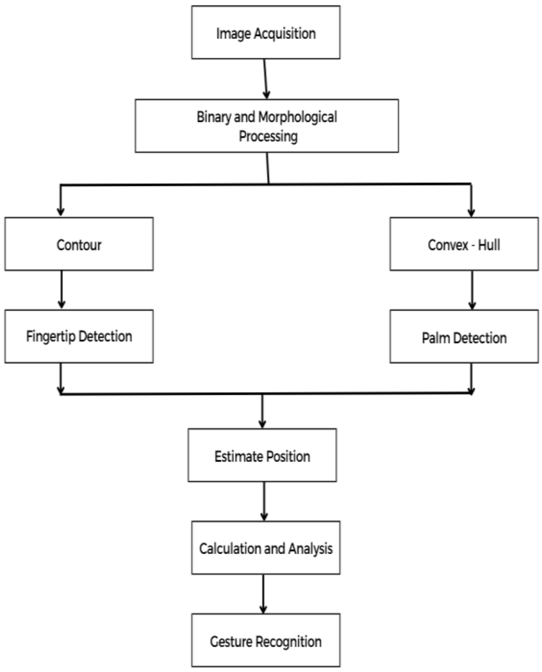


Fig.6: Workflow for Hand Detection

- **User Interface (UI):**

\- The User Interface of this web application is user-friendly, flexible and it can also be translated into **16 different languages** including **Indian** and some **foreign languages.**

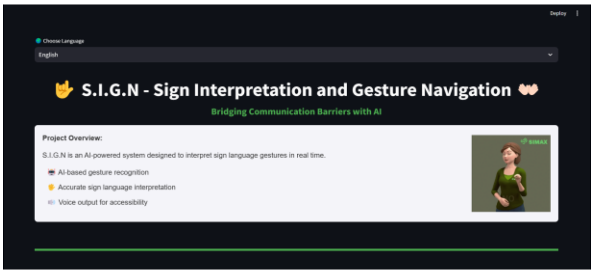


Fig.7: UI Page with various language translation

Fig.8: UI (background)

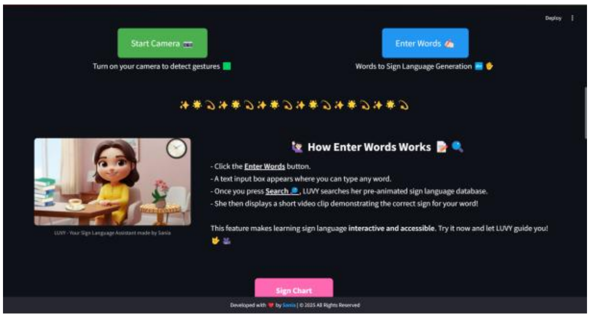


Fig.9: UI (phases and working)

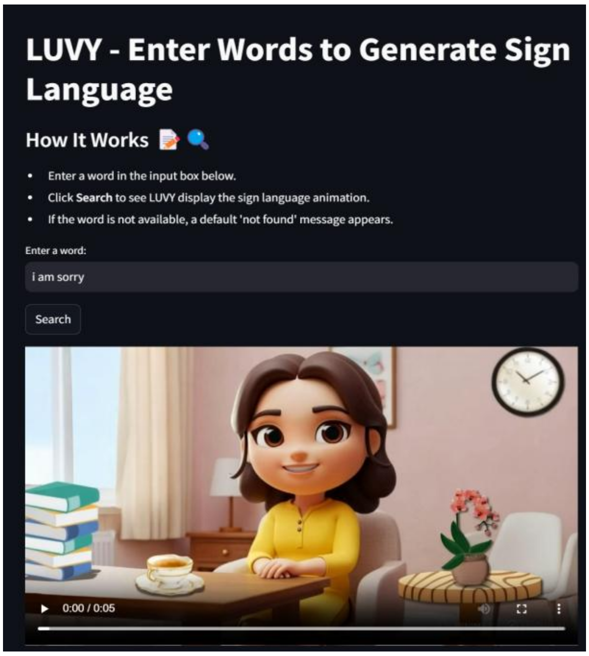


Fig.10: LUVY - Our Sign Language Assistant

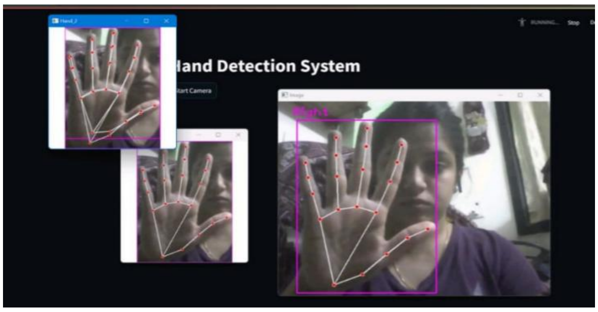


Fig.11: Hand & Fingertips landmarks

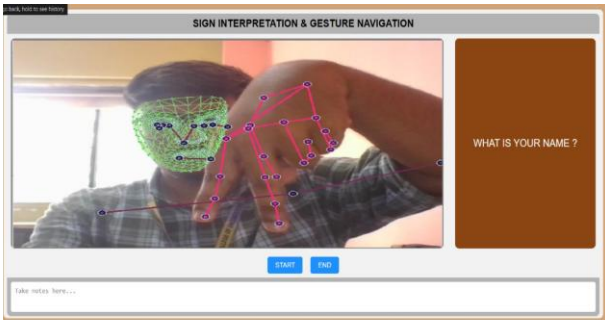


Fig.12: Results (Sentence Formation)

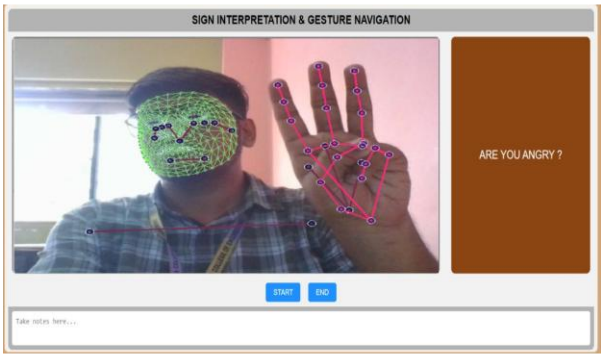


Fig.13: Results (Sentence Formation)

Fig.14: Results (Accuracy)

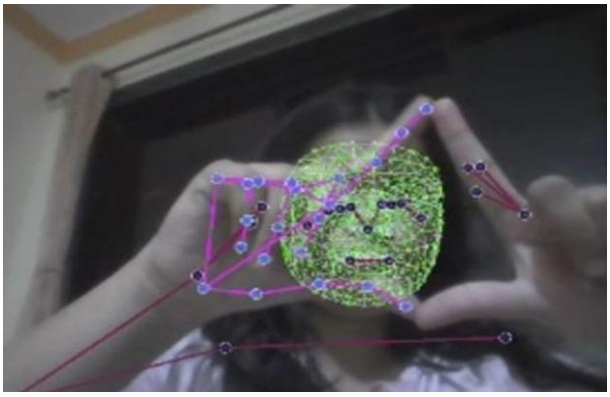


Fig.15: Double Hand Detection

VII. TESTING & RESULT ANALYSIS

1. **Unit Testing**

Objective: To test individual components of the system, such as gesture recognition, hand detection, and text conversion modules, in isolation.

2. **Integration Testing**

Objective: To verify that all modules function correctly when combined.

Test Case	Description	Expected Outcome	Status
Hand detection	Detects hand landmarks	Hand is successfully detected	Pass
Gesture recognition	Recognize specific signs	Correct text output is displayed	Pass
Lighting condition	Tests under different lighting	Recognition remains accurate	Pass
User Interface	UI is user-friendly and accessible	Users can interact easily	Pass

3. **User Acceptance Testing (UAT)**

Objective: To ensure the system meets user requirements and is easy to use.

VIII. FUTURE PROSPECT

8.1. AI-Driven Real-Time Sign Language Interpretation

Future AI models will better understand regional sign languages, facial expressions, and contextual meaning. Smart glasses or AR devices could provide instant sign language translation for students in real-world learning scenarios. AI could integrate sign language with text, speech, and haptic feedback for a richer learning experience.

8.2 Gesture-Based Learning Expanding Across Subjects

Gesture-controlled virtual labs will allow students to conduct physics and chemistry experiments without physical tools. Gesture-based tools can help students learn sign language interactively, bridging the gap between hearing and non-hearing students.

8.3 Integration with Virtual Reality (VR) and Augmented Reality (AR)

VR and AR classrooms could allow students to interact with educational content using sign language and gestures. Gesture-controlled VR classrooms will enable students to participate in global educational programs seamlessly. AI will analyze student gestures and interactions to tailor learning experiences.

8.4 Smart Classrooms with AI & IoT Integration

Future classrooms will allow students to control digital screens, smartboards, and e-books using gestures. Cameras and sensors could analyze students' gestures to measure engagement and adapt lessons accordingly.

IX. CONCLUSION

This paper presents a Real-time Sign Language Interpreter for Indian Sign Language (ISL) and a keyword to sign generator that leverages computer vision and rule-based processing techniques to bridge communication barriers for the hearing-impaired community. By utilizing OpenCV, MediaPipe, CNN and predefined gesture recognition methods, the system provides a cost-effective and efficient alternative to human interpreters. Future developments can expand functionality to support dynamic gestures and improve adaptability across diverse environments, promoting greater inclusivity and accessibility.

X. REFERENCES

[1] Kandasamy S; Mowlieshwaran S; Kiran R; Kishore DS; Karthikeyan M "Sign Language Recognition using Python and OpenCV", pp. 88 - 92, DOI: 10.1109/ICSMDI57622.2023.00023

[2] Rashmi Gaikwad and Lalita Admuthe; "Real-Time Sign Language Recognition of Words and Sentence Generation using MediaPipe and LSTM", pp. 287 - 297, DOI:10.1007/978-981-97-0180-3_23

[3] Aakash Deep; Aashutosh Litoriya; Akshay Ingole; Vaibhav Asare; Shubham M Bhole; Shantanu Pathak, "Realtime Sign Language Detection and Recognition", 2022, DOI: 10.1109/ASIANCON55314.2022.9908995

[4] Vishwa Hariharan Iyer; U.M Prakash; Aashrut Vijay; P. Sathishkumar "Sign Language Detection using Action Recognition", 2022. DOI: 10.1109/ICACITE53722.2022.9823484

[5] A. Moryossef, I. Tsochantaridis, J. Dinn, N.C. Camgoz, R. Bowden, T. Jiang, "Evaluating the immediate applicability of pose estimation for sign language recognition", pp. 3434-3440, DOI:

[6] K. Amrutha, P. Prabu and J. Paulose, "Human Body Pose Estimation and Applications", 2021, pp. 1-6, DOI: 10.1109/i-PACT52855.2021.9696513

[7]Pradnya D. Bormane, S. D. Shirbahadurkar, "Indian Sign Language Recognition: Support Vector Machine Approach", DOI: 10.52783/anvi.v27.1438

[8] A. Wadhawan and P. Kumar, "Deep learning - based sign language recognition system for static signs", Neural computing and applications, 2020, DOI: 10.1007/s00521-019-04691-y

[9] S. Katoch, V. Singh, & U. Tiwary, "Indian sign language recognition system using surf with svm and cnn", Array, vol. 14, p. 100141, 2022.

[10] L. Chandwani, "Gesture based sign language recognition system using mediapipe", 2023. ISSN: 1092-910X

[11] D, Karthika Renuka and Ashok Kumar L, "Indian Sign Language Recognition Using Deep Learning Techniques",(2022). DOI: 10.34256/ijcci2214

[12] Poornima, B.V and S. Srinath, "Frequency and Spatial Domain-Based Approaches for Recognition of Indian Sign Language Gestures," (2024), pp. 660 - 669, DOI: 10.17485/IJST/v17i7.2836