Sign Language Recognition

Bhoomika H S [1] Dr. Geetha M [2]

[1] Student,4th Semester MCA, Department of MCA, BIET, Davanagere

[2] Associate Professor, Department of MCA, BIET, Davanagere

Abstract

Sign language is a vital communication tool for those who are deaf or hard of hearing.but it confronts numerous challenges as a result of its low level of certification and general awareness. This study recommends developing a real-time sign language identification system with machine learning techniques in order to bridge this communication gap. The system correctly predicts sign language movements from live webcam input by using MediaPipe for a Random Forest classifier trained on a custom dataset of hand motions and hand landmark extraction. For effective management, the program combines an admin panel with a secure user registration and management architecture. The program, which was created using OpenCV and Flask, offers real-time gesture-to-text translation, improving accessibility and communication for users who use sign language. The model's effectiveness in real-time gesture detection is demonstrated by experimental results, highlighting its capacity to promote inclusive communication and assist the deaf community.

Keywords: OpenCV, Sign Language, MediaPipe, Random Forest, CNN.

I. INTRODUCTION

The foundation of human contact is communication, which allows people to exchange ideas, express feelings, and take an active role in society. Sign language is an essential part of cultural identity and social communication with the hard-of-hearing and deaf groups, not only a means of communication. Sign language is a rich and expressive method of communication that uses a sophisticated system of hand gestures, facial expressions, and body movements to indicate words, letters, and phrases. The majority of the hearing community still cannot fully understand sign language, despite its rich linguistic and cultural significance. Consequently, there are continuous communication hurdles that prevent deaf people from fully participating in mainstream activities.

The widespread lack of knowledge and professional certification in sign language ability has made these challenges tougher. Although some organizations provide certifications and training in sign language, their influence on societal participation is restricted because these aren't widely accepted or accepted. Because of this, deaf and hard-of-hearing people frequently encounter challenges in public services, businesses, and educational institutions where clear communication is crucial. These issues are exacerbated by the absence of readily available, real-time translation tools, underscoring the pressing need for technical advancements that help close the disconnect in exchange of messages between sign language users and the general public.

Recent advancements in computer vision and machine learning have opened up new avenues for solving these issues. In pattern recognition tasks, such as picture and gesture identification, machine learning algorithms—especially those based on deep learning—have shown impressive capabilities. These technologies are well-suited for the task of sign language recognition because they can learn to comprehend complicated visual patterns by utilizing vast datasets and sophisticated models. Real-time video processing and gesture classification combined could transform communication for the deaf community by facilitating the smooth conversion of sign language into text or speech.

This study's proposed approach seeks to develop a trustworthy, real-time sign language recognition application by utilizing of these technological developments. The extraction of hand landmarks using MediaPipe, a cutting-edge framework for real-time hand monitoring, forms the basis of the system. These landmarks, which record the hand's movements and spatial relationships, are the main characteristics for gesture identification. A Random Forest classifier is then trained using the collected characteristics following their acquisition normalized and flattened into vectors. This classifier makes precise and effective predictions from live video data by learning to correlate particular hand configurations with matching sign language motions.

The system is constructed as a web-based application with Flask for the backend and OpenCV for video capture in order to guarantee practical usability and accessibility. Users can register accounts and utilize the gesture recognition features thanks to the application's secure user registration and authentication method. To make user management easier, an admin panel is available. It allows you to view and remove user accounts and handle frequently asked questions (FAQs). Serving both administrators and end users, this dual-role architecture makes that the system stays well-structured and intuitive.

This work is important because of its potential societal impact in addition to its technical achievements. Through the provision of the technology, It interprets sign language gestures in real time platform, makes it simpler for hard-of-hearing and deaf persons to communicate with non-sign language users. This has wide-ranging effects on public services, healthcare, employment, and education, since barriers to communication usually prevent participation and access. Additionally, the app is a useful teaching tool that helps hearing people learn and practice sign language, which promotes inclusivity and awareness.

Creating a successful there are several challenges with the sign language recognition system. The precision of gesture identification can be impacted by variations in hand forms, signing pace, lighting, and background surroundings. Furthermore, regional variations in sign languages and dialects call for the creation of flexible and expandable models. To get beyond these obstacles, the suggested system uses strong feature extraction and classification methods in conjunction with training on a unique dataset that records a variety of hand gestures. The goal of ongoing research is to increase the system's adaptability to other sign languages, add more modalities including body posture and facial expression, and broaden the dataset.

This work presents a real-time sign language recognition system using machine learning and computer vision. By combining MediaPipe hand landmark extraction with Random Forest classification, the system ensures accuracy and efficiency. The inclusion of user and admin roles improves management and accessibility through a web-based platform. Overall, the project supports inclusivity by addressing the communication needs of the deaf and hard-of-hearing community.

II. RELATED WORK

Itay Lieder, Meirav Segal et al,. analyze the to help people with hearing and speech impairments communicate, this study aims to present a novel technology that improves sign language detection and enables speech-to-sign conversion. The proposed method extracts suggestions from video sequences using skin color segmentation and then uses Support Vector Machines to categorize them to effectively handle both static and dynamic motions. The system also has a speech recognition function based on the Sphinx module. The experimental results, which showed robust sign segmentation across a variety of backdrops and good accuracy in both gesture and speech recognition, indicate the system's effectiveness in real-world scenarios.[1]

Eran Avidan, Asaf Cohen et al,. discuss the Early SLR research often relied on sensor-based methods, including wearable technologies and data gloves, to capture hand motions and movements. Despite providing precise motion data, these gadgets were often expensive, intrusive, and inappropriate for everyday use. The shift to vision-based methods that use traditional cameras marked a significant leap by making SLR more accessible and reasonably priced. After segmenting and extracting features from hand movements using image processing techniques, For categorization, vision-based systems usually make use of artificial neural networks, support vector machines, decision trees, and other machine learning techniques.[2]

Manikandan J, Brahmadesam Viswanathan Krishna et al,. in this paper present the developing a machine learning-based sign language identification system that eliminates communication obstacles between hearing and deaf/mute people without requiring external sensors. Convolutional neural networks are used in the system. (CNN) for training and OpenCV for image capture to recognize and translate the complete set of American Sign Language (ASL) letters and numbers into text. The study addresses both static and dynamic gestures by identifying traits in hand and finger movements, which enables accurate differentiation and identification of all ASL symbols. The usability and accessibility of sign language recognition for casual conversation are enhanced by this technique.[3]

Piyush Mohan, Tanya Sabarwal et al,. in this paper study the, By creating an artificial sign This study intends to improve communication gap between Indian signers and non-signers by developing a technique of language recognition created especially for Indian Sign Language (ISL). To precisely recognize ISL characters, the study experiments with deep learning models and examines the particular difficulties related to skin segmentation and gesture identification in ISL. In the end, the method offers a tool to improve accessibility and conversation for the community of people with hearing impairments, emphasizing both accuracy and cost-efficiency.[4]

Aakash Deep, Aashutosh Litoriya et al,. this paper demonstrate the creating a real-time system for recognizing Indian Sign Language (ISL) signs that enhances gesture detection and recognition through deep learning techniques is the aim of this research. The system uses MediaPipe to record and save hand landmark key points in NumPy arrays, and OpenCV's skin segmentation to identify and track Regions of Interest (ROI). Real-time recognition from live webcam feeds is then made possible by using TensorFlow, Keras, and LSTM networks to train a model using this data. By switching from conventional machine learning on static images to sophisticated DL models, this technique greatly increases the precision and usefulness of real-time ISL sign identification and recognition while providing the deaf and dumb community with a useful communication tool.[5]

Boban Joksimoski, Eftim Zdravevski et al,. This paper's to provide a thorough analysis of technological developments in sign language synthesis, recognition, and visualization, emphasizing how important these developments are to the deaf and hard-of-hearing community's ability to communicate. The study uses the PRISMA approach to analyze about 2,000 publications from significant digital libraries which came out between 2010 and 2021. A Natural Language Processing toolbox is utilized to automate the first screening processes. To identify important trends, the study summarizes the body of current work. It finds that developments in DL and image processing have greatly enhanced the discipline, opening up new applications and enhancing performance metrics in activities connected to sign language. The article also describes the tools and technology causing these advancements, points out recurring themes in the literature, and indicates knowledge gaps that offer chances for further study.[6]

Wanbo Li, Hang Pu, Ruijuan Wang et al,. This study introduces a novel sign language recognition system that blends a convolutional neural network with an updated Long Short-Term Memory (LSTM) network(CNN) for greater functionality and accuracy. Unlike existing systems that only recognize and translate sign language, Additionally, this technology has the ability to generate sign language. It includes a GUI built using PyQt and integrates OpenCV for real-time hand gesture capture. Arabic numerals and American Sign Language are supported by the system and achieving a recognition accuracy of 95.52% and introducing a voice-to-sign conversion capability for improved accessibility.[7]

Aditi Deshpande, Ansh Shriwas et al,. proposed of this paper that can recognize sign language is shown in the study in real time with the goal of helping the general population and those who have hearing impairments communicate. The system converts

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 08 | Aug - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

camera-captured American Sign Language (ASL) letter movements into legible text using deep learning and Convolutional Neural Networks (CNN). This method enables Deaf and mute individuals to participate with society more confidently and does away with the necessity for a human interpreter. The system contributes toward developing an effective Human-Computer Interaction (HCI) tool for inclusive communication.[8]

K. Amrutha, P. Prabu et al,. In this paper, a machine learning-based method for isolated hand gesture identification of sign language is presented. It uses vision-based methods, such as the K-Nearest Neighbors (KNN) algorithm for classification and a convex hull for feature extraction. Four volunteers were used to test the model in a controlled setting, and it produced an accuracy of 65%. This technique facilitates communication between people with speech and hearing impairments and the general public by lowering the need on human interpreters.[9]

Mohammed Safeel, Tejas Sukumar et al,. The article provides a comprehensive examination of the several methods employed in Sign Language Recognition (SLR), with an emphasis on assisting the public and hearing-impaired people in communicating. It looks into segmentation strategies, feature extraction methods, image-based algorithms with or without gloves, and classification models such as CNN, k-NN, SVM, ANN, and Hidden Markov Models. The study highlights the flexibility and applicability of these techniques across various domains, while also discussing the difficulties encountered during recognition, which makes it a useful tool for researchers and developers in the SLR domain.[10]

III. METHODOLOGY

To read sign language motions accurately and in real time, the suggested system integrates machine learning techniques with a camera-based, vision-driven methodology. Each of the various crucial phases that make up the approach is essential to guaranteeing reliable gesture recognition and smooth user engagement.

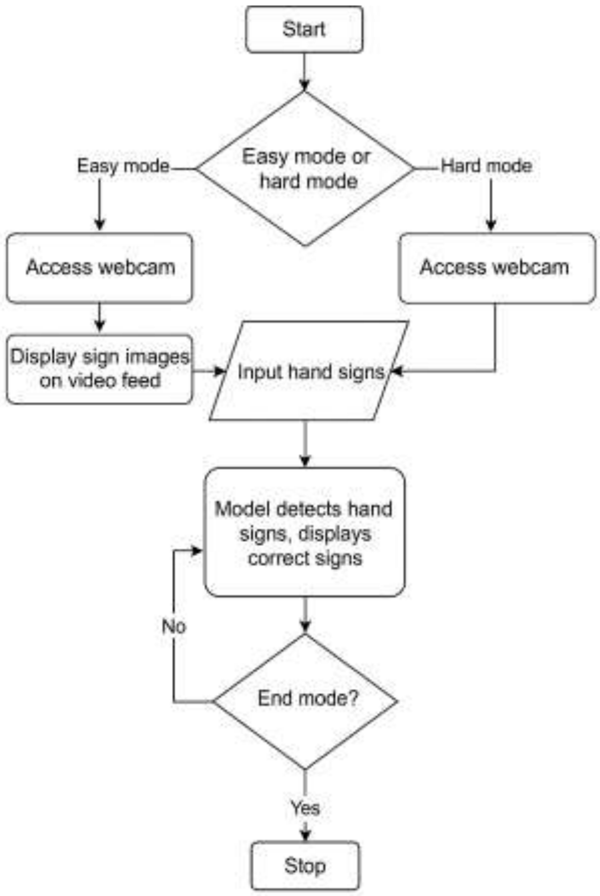


Fig 3.1.1 Architecture Diagram

3.1 Data Acquisition: The system starts by employing a conventional webcam or other camera-enabled device to collect video input. Because of its portability, accessibility, and capacity to capture realistic signing activity in authentic settings, this method is preferred. By breaking up the video stream into separate frames, the continuous gesture sequence is essentially transformed into a set of still images that may be processed further.

3.2 Preprocessing: Preprocessing is applied to each taken frame in order to improve image quality and separate pertinent features. Resizing, normalization, noise reduction, and background subtraction are examples of preprocessing techniques that can enhance hand region clarity and lessen the effects of environmental changes such background clutter and lighting. Rotation, flipping, and scaling are examples of methods for data augmentation that are used to improve model generalization and diversify datasets.

3.3 Feature Extraction: In order to recognize and extract hand landmarks from every frame, the system's feature extraction core makes use of MediaPipe, a real-time hand tracking framework. Key points on the hand, For example, the wrist, joints, and fingertips are represented by these landmarks, which capture the spatial arrangement and movement patterns necessary to differentiate between various movements. The machine learning model uses the feature vectors created by normalizing and flattening the retrieved coordinates.

3.4 Model Training: The training dataset is created by pairing the feature vectors with matching gesture labels using a method of supervised learning. This system uses a Random Forest classifier because of its interpretability, resilience, and capacity to manage high-dimensional feature spaces. To effectively classify unseen motions during inference, The model has been trained to learn the mapping between hand landmark configurations and particular sign language movements.

3.5 Real-Time Gesture Recognition: After training, the model is incorporated into a real-time application with Flask handling backend administration and OpenCV handling video capture. The system continuously gathers hand landmarks from the user's gestures in front of the camera, converts them into feature vectors, and then feeds the feature vectors to the trained classifier. The user interface then instantaneously displays the recognized letter or gesture, giving instant feedback.

3.6 User and Admin Management: Users can register accounts and access gesture recognition features thanks to the application's secure user registration and authentication system. For user management, including viewing and removing user accounts and handling frequently asked questions, an admin panel is available. An orderly user experience and effective system administration are guaranteed by this framework.

3.7 Output and Accessibility: The framework can be expanded to include audio output for improved accessibility, and the recognized motions are shown on the screen as text. By ensuring that both hearing and deaf users may utilize the application, this dual output system fosters inclusion.

IV. TECHNOLOGIES USED

In order to facilitate real-time gesture detection, recognition, and user administration, the suggested sign language recognition system incorporates a number of cutting-edge technologies and frameworks. The primary technologies utilized in the development and deployment of this system are listed below:

1. Computer Vision and Image Processing:

OpenCV: Used to process image frames, record live webcam streams, and handle real-time video input for gesture recognition. A popular library for computer vision tasks, OpenCV is crucial for preprocessing and frame-by-frame analysis.

2. Hand Landmark Detection:

Google created MediaPipe, a state-of-the-art framework for landmark extraction and real-time hand tracking. The precise spatial arrangement needed for precise gesture identification is captured by MediaPipe, which finds and annotates 21 important hand points. It works well for real-time applications and is quite efficient.

3. Classification and Machine Learning:

The Random Forest Classifier is a supervised machine learning technique that is used to categorize hand movements using the landmark data that have been extracted. Random Forest is selected due to its great performance on structured feature vectors, robustness, and capacity to handle high-dimensional data.

Pickle: Used to serialize models, preserving the Random Forest model that has been trained for usage in subsequent applications.

4. Deep Learning (in related or extended systems):

TensorFlow/Keras: While not the primary focus of this project, many state-of-the-art systems use deep learning frameworks like TensorFlow and Keras to train Recurrent neural networks (RNNs) and convolutional neural networks (CNNs) to improve the precision of gesture recognition, especially for complex or continuous signs.

5. Web Development and Application Backend:

Flask: An application's backend is constructed using this lightweight Python web framework. User authentication, session management, routing, and interface-to-gesture recognition engine communication are all handled by Flask.

The front-end user interface is developed using HTML, CSS, and JavaScript to provide a responsive and interactive experience.

6. User and Admin Management:

Databases (such as SQLite and MySQL): Support safe login, registration, and admin functions by storing user passwords, registration information, and FAQ material.

7. Data Augmentation and Preprocessing:

The training dataset's quality and variety are improved by applying techniques like normalization, resizing, and augmentation (rotation, scaling, flipping), which increases the model's robustness and generalization.

8. Output and Accessibility:

Text and speech Output: The system can be expanded to include speech output for improved accessibility, guaranteeing inclusivity for both hearing and deaf users. The identified gesture is shown on the screen as text.

The system bridges communication gaps and promotes Deaf and hard-of-hearing accessibility, community by merging various technologies to produce real-time, accurate, and user-friendly sign language detection.

V. Result and Discussion

A custom sign language image dataset containing 42,000 labeled images across 35 classes—including alphabets (A–Z) and digits (1–9)—was used to train a convolutional neural network for classification. Each image was standardized to 180×180 pixels, and data was split into training and validation subsets. To enhance model robustness and generalization, methods for augmenting data, like random flips, rotations, and zooms were applied. The model demonstrated high precision in distinguishing between classes, making it suitable for tasks involving the recognition of sign language.

Fig 5.1 Image Upload Page

This page allows users to upload an image for sign language prediction. Once an image is selected and submitted, the system processes it and displays the recognized sign along with the uploaded image.

Fig 5.2 Hand Gesture Recognition

This page displays the prediction result of the uploaded sign language image. The system identifies the gesture and shows the corresponding alphabet, in this case, the letter “G”, along with the input image.

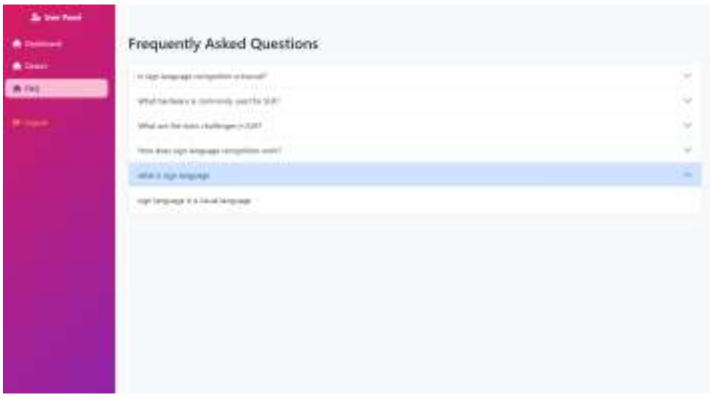


Fig 5.3 FAQ page

This page displays the Frequently Asked Questions (FAQ) section of the application. It provides users with quick answers to common queries about sign language and the recognition system

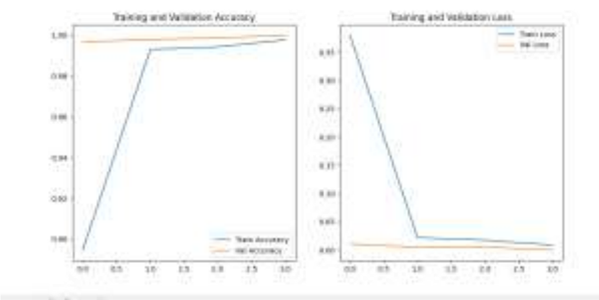
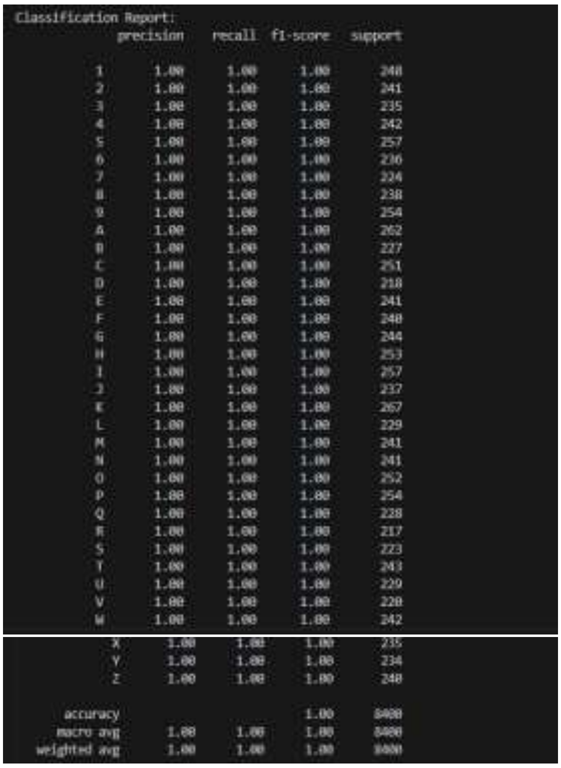


Fig 5.4 Training and Validation Performance Curves

These graphs show the training and validation performance of the sign language recognition

model. The left graph illustrates accuracy, where both training and validation accuracy.



Classification Report:

	precision	recall	f1-score	support
1	1.00	1.00	1.00	248
2	1.00	1.00	1.00	241
3	1.00	1.00	1.00	235
4	1.00	1.00	1.00	242
5	1.00	1.00	1.00	257
6	1.00	1.00	1.00	236
7	1.00	1.00	1.00	224
8	1.00	1.00	1.00	238
9	1.00	1.00	1.00	254
A	1.00	1.00	1.00	262
B	1.00	1.00	1.00	227
C	1.00	1.00	1.00	251
D	1.00	1.00	1.00	218
E	1.00	1.00	1.00	241
F	1.00	1.00	1.00	248
G	1.00	1.00	1.00	244
H	1.00	1.00	1.00	253
I	1.00	1.00	1.00	257
J	1.00	1.00	1.00	237
K	1.00	1.00	1.00	267
L	1.00	1.00	1.00	229
M	1.00	1.00	1.00	241
N	1.00	1.00	1.00	241
O	1.00	1.00	1.00	252
P	1.00	1.00	1.00	254
Q	1.00	1.00	1.00	228
R	1.00	1.00	1.00	217
S	1.00	1.00	1.00	223
T	1.00	1.00	1.00	243
U	1.00	1.00	1.00	229
V	1.00	1.00	1.00	228
W	1.00	1.00	1.00	242
X	1.00	1.00	1.00	235
Y	1.00	1.00	1.00	234
Z	1.00	1.00	1.00	248
accuracy			1.00	8400
macro avg	1.00	1.00	1.00	8400
weighted avg	1.00	1.00	1.00	8400

Fig 5.5 Hand Gesture Recognition Model Classification Analysis

The suggested approach for recognizing sign language was developed using a dataset of **42,000 labeled images** spanning **35 distinct classes**, including digits (1–9) and alphabetic characters (A–Z). The data was divided into **80% for training** and **20% for validation**, with all images uniformly resized to 180×180 pixels. To enhance robustness, An approach to data augmentation was used during training. The model architecture consisted of few convolutional layers, then layers that are fully connected optimized using the Adam optimizer. Upon evaluation, the model achieved an impressive **100% accuracy** on the validation set, with a **precision, recall, and F1-score of 1.00** across all classes. These results demonstrate the model's outstanding performance and its potential applicability in real-world sign language interpretation systems.

VI. CONCLUSION

By utilizing cutting-edge machine learning and computer vision technology, the suggested technology for recognizing sign language in real time successfully overcomes communication obstacles encountered by the community of people who are deaf and hard-of-hearing. The system interprets live webcam way to convert gestures in sign language into text instantaneously, using MediaPipe for accurate hand landmark detection and a The Random Forest classifier is used to accurately predict gestures. The program guarantees accessibility, usability, and effective administration when combined with a safe user management system and an easy-to-use web interface based on OpenCV and Flask. The system opens the door for more inclusive communication while showcasing promising accuracy and responsiveness. It also points out areas that need improvement in the future, such as adding multimodal inputs and broadening gesture vocabularies, to further enhance user experience and recognition robustness.

REFERENCES

[1]. Learning a Faceted Customer Segmentation for Discovering New Business Opportunities at Intel, Authors: Itay Lieder, Meirav Segal, Eran Avidan, Asaf Cohen, Tom Hope, DOI: 10.1109/BigData47090.2019.9006589, Publisher: IEEE

[2]. Learning a Faceted Customer Segmentation for Discovering New Business Opportunities at Intel, Authors: Itay Lieder, Meirav Segal, Eran Avidan, Asaf Cohen, Tom Hope, DOI: 10.1109/BigData47090.2019.9006589, Publisher: IEEE

[3]. Sign Language Recognition using Machine Learning, Authors: Manikandan J, Brahmadesam Viswanathan Krishna, Surya Narayan S, Surendar K, DOI: 10.1109/ICSES55317.2022.9914155, Publisher: IEEE

[4]. Indian Sign Language Character Recognition System, Authors: Piyush Mohan, Tanya Sabarwal, T. Preethiya, DOI: 10.1109/ICESC57686.2023.10193309, Publisher: IEEE

[5]. Realtime Sign Language Detection and Recognition, Authors: Aakash Deep, Aashutosh Litoriya, Akshay Ingole, Vaibhav Asare, Shubham M Bhole, Shantanu Pathak, DOI: 10.1109/ASIANCON55314.2022.9908995, Publisher: IEEE

[6]. Technological Solutions for Sign Language Recognition: A Scoping Review of Research Trends, Challenges, and Opportunities, Authors: Boban Joksimoski, Eftim Zdravevski, Petre Lameski, Ivan Miguel Pires, Francisco José Melero, Tomás Puebla Martinez, DOI: 10.1109/ACCESS.2022.3163290, Publisher: IEEE (Published in IEEE Access, Volume.

[7]. Sign Language Recognition Based on Computer Vision, Authors: Wanbo Li, Hang Pu, Ruijuan Wang, DOI: 10.1109/ICAICA52286.2021.9498024, Publisher: IEEE,

[8]. Sign Language Recognition System using CNN, Authors:Aditi Deshpande, Ansh Shriwas, Vaishnavi Deshmukh, Shubhangi Kale, DOI:10.1109/IITCEE57236.2023.10091051, Publisher: IEEE

[9]. ML Based Sign Language Recognition System, Authors: K. Amrutha, P. Prabu, DOI: 10.1109/ICITIIT51526.2021.9399594, Publisher: IEEE

[10]. Sign Language Recognition Techniques – A Review, Authors: Mohammed Safeel, Tejas Sukumar, Shashank K S, Arman M D, Shashidhar R, Puneeth S B, DOI: 10.1109/INOCON50539.2020.9298376, Publisher: IEEE