

SIGN LANGUAGE RECOGNITION SYSTEM FOR SEAMLESS HUMAN-AI INTERACTION

Harshini P. J ¹, Mohammed Atheeque Rahman ²,
James Allen Raj ³, Dr.P.Durgadevi ⁴

1,2,3,4 Department of CSE,

*SRM Institute of Science and Technology, Vadapalani campus,
Chennai.*

ABSTRACT:

Communication barriers persist for individuals who use sign language as their primary mode of interaction, particularly in digital environments where conversational AI bots are increasingly prevalent. This paper presents a novel approach to addressing these barriers through the development of a Sign Language Recognition System (SLRS) integrated with conversational AI capabilities. The SLRS utilizes advanced computer vision techniques and machine learning algorithms, specifically employing a Random Forest Classifier, achieving an exceptional accuracy of 0.9961. Additionally, a comparison between CNN, KNN, and Random Forest classifiers underscores the superiority of Random Forest in accuracy and performance. Integration with the conversational AI bot "Gopal," powered by the Google Gemini Model, facilitates natural language interactions, allowing users who use sign language to communicate seamlessly with digital systems. Evaluation of the SLRS demonstrates promising performance in accurately recognizing a diverse range of sign language gestures and generating contextually relevant responses. User testing sessions highlight the system's responsiveness and effectiveness in bridging the communication gap between users who use sign language and conversational AI bots. The development of the SLRS represents a significant step towards promoting inclusivity and accessibility in digital communication for those who communicate with sign language. By enabling seamless interaction with conversational AI bots, the SLRS empowers users who use sign language to engage more effectively with digital interfaces and access information effortlessly. Continued research and refinement of the SLRS hold the potential to further enhance communication accessibility and foster greater inclusion in digital environments.

Keywords: *Sign Language Recognition, Human-AI Interaction, Object Detection, Machine Learning, Conversational AI.*

1. INTRODUCTION:

Sign language serves as a crucial means of communication for individuals with speaking impairments, offering them the ability to express themselves and interact with the world around them. Despite its importance, the communication barrier between those who use sign language and those who don't remains a significant challenge, impacting various aspects of daily life such as education, employment, and social interaction. This challenge stems from the inherent differences between sign language and spoken language, as well as the limited accessibility of tools and technologies tailored to the needs of users who use sign language .

In response to these challenges, this research project endeavors to develop and evaluate a novel Sign Language Recognition System (SLRS) that aims to bridge the communication gap between those who use sign language and those who don't. Real-time recognition and interpretation of sign language motions is made possible by the SLRS, which facilitates smooth communication between users who use sign language and conversational agents driven by artificial intelligence. By leveraging advancements in computer vision and machine learning [1], the SLRS seeks to empower individuals with speaking impairments to communicate effectively and access information and services with ease.

At the heart of the SLRS is its ability to precisely identify and interpret motions in sign language as text and speech outputs [2]. This functionality is achieved through a combination of sophisticated algorithms for hand gesture recognition [3], feature extraction, and machine learning-based classification. By examining the temporal and spatial properties of hand movements, the SLRS can distinguish between different sign language gestures and generate corresponding textual and auditory representations [4]. This capability not only facilitates communication between users who use sign language and AI systems but also enhances the accessibility and usability of digital interfaces for individuals with speaking impairments.

One of the key innovations of this project lies in the incorporation of the SLRS with AI-powered conversational agents, commonly known as chatbots. This integration brings forth several significant contributions:

Integration with Chatbots: The project integrates the Sign Language Recognition System (SLRS) with AI-powered conversational agents, facilitating seamless interaction between users who use sign language and AI systems.

Natural Language Understanding and Generation: Through this integration, the SLRS gains the ability to interpret and respond to sign language inputs by including natural language creation and comprehension skills.

Sign Language Interaction: Users can now interact with AI systems using sign language inputs, breaking down communication barriers and enhancing accessibility.

Multimodal Response: The SLRS provides responses in both textual and auditory formats, catering to diverse user preferences and accessibility needs.

Efficiency and Naturalness: The integration enhances the efficiency and naturalness of human-AI interaction, making communication more intuitive and effective.

Accessibility and Inclusivity: By opening up new possibilities for accessibility and inclusivity in digital communication, the project empowers users who use sign language to participate more fully in society.

Through seamless integration and intelligent interaction, the SLRS aims to empower users who use sign language and reduce the communication gap that exists between sign language and spoken language users.

In the following sections of this paper, Before presenting the methodology, we provide a thorough analysis of the existing literature in the discipline of SLR, or sign language recognition (**Section 2: Literature Review**), encompassing various studies that delve into deep learning models, training techniques, and innovative systems tailored for sign language communication and recognition. we will delve deeper into the methodology employed in developing the SLRS, including data collection, hand gesture recognition, machine learning model training, and integration with conversational AI (**Section 3: Methodology**). We will also discuss the architecture and implementation details of the SLRS (**Section 4: Architecture and Implementation**), present the results of system evaluation and user testing (**Section 5: Results**), and conclude with insights into future research directions and potential applications of the SLRS in real-world settings (**Section 6: Discussion and Future Directions**). Through rigorous research and innovation, we aim to contribute to the advancement of accessibility and inclusion for individuals with speaking impairments, fostering a more inclusive and equitable society for all.

2. LITERATURE REVIEW:

The research paper authored by Abu Saleh Musa Miah, Md. Al Mehedi Hasan, Satoshi Nishimura, and Jungpil Shin delves into deep learning models and training techniques customized for Sign Language Recognition (SLR) [2]. Their investigation encompasses diverse datasets, including a novel Greek Sign Language (GSL) dataset, aiming to address key challenges in SLR such as accurate segmentation and sequence modeling. A notable aspect of their work involves evaluating variations of the Connectionist Temporal Classification (CTC) technique, namely EnCTC and StimCTC, which effectively handle issues like ambiguous boundaries between sign gestures and intra-gloss dependencies.

The research paper authored by Ajay S, A. Potluri, S. M. George, G. R, and A. S introduces a system utilizing sensor gloves to detect Indian Sign Language (ISL) gestures and convert them into audible speech [5]. The gloves are equipped with a range of sensors and modules, including Bluetooth, RF, Inertial Measurement Units (IMU), flex, touch, and Arduino Nano microcontrollers. The integration of sensor data

with machine learning classification algorithms enhances gesture recognition accuracy, addressing the communication gap for the speech-impaired population in India.

Another paper by M. Al-Hammadi, G. Muhammad, W. Abdul, M. Alsulaiman, M. A. Bencherif, T. S. Alrayes, H. Mathkour, and M. A. Mektiche presents a novel system for dynamic hand gesture recognition employing deep learning architectures [3]. They address obstacles at hand identification of gestures such as segmentation and sequence modeling, offering a comprehensive solution that leverages both local and global features [6]. Experimentation on diverse datasets showcases the system's superior performance compared to existing approaches, demonstrating its potential for real-world applications in translating into sign language and human-computer interaction.

Furthermore, Ben Atitallah et al. introduce an Electrical Impedance Tomography (EIT) imaging system tailored for hand sign recognition and monitoring [7]. Their system, utilizing a simple setup with eight electrodes, achieves high classification accuracy for American Sign Language numbers from 0 to 9. Notably, the Convolutional Neural Network (CNN) classifier outperforms Support Vector Machine (SVM) and Softmax classifiers, showcasing robustness across different subjects with only a slight decrease in accuracy.

Researchers from the College of Engineering Pune present a pioneering study focused on bridging communication gaps for the speechless community [8]. Their novel system utilizes KNN, SVM, and CNN algorithms for the identification of hand gestures, highlighting the potential of machine learning in enhancing inclusivity and accessibility in communication.

Additionally, researchers focus on utilizing Convolutional Neural Networks (CNNs) to recognize dynamic signs in Indian Sign Language (ISL), aiming to bridge communication gaps for the hearing-impaired community [9]. The study emphasizes the possibilities of deep learning methods for real-time interaction and human-computer interfaces.

Another study explores the creation of a prototype for a sign language interfacing system, leveraging advancements in virtual reality and gesture recognition technology to enhance communication and interaction for those who have hearing loss [10].

Muhammad Abid, Emil Petriu, Ehsan Amjadian, and Domenico Grimaldi focus on dynamic sign language recognition tailored for smart home applications [11], highlighting the effectiveness of their system for smart home integration.

Furthermore, Aparna Mohanty, Sai Saketh, Rajiv Sahay, and S Rambhatla Sahay investigate deep learning CNN techniques for robust hand gesture recognition [12], addressing challenges such as cluttered backgrounds. Despite certain limitations, their proposed CNN model achieves high accuracy and performance, suitable for applications in sign language recognition and human-computer interaction.

Lastly, Muhammad Al-Qurishi and Riad Souissi conduct a comprehensive literature survey from 2014 to 2021 [13], evaluating 84 relevant works and scrutinizing methodologies from traditional Hidden Markov Models to modern deep learning techniques. They highlight challenges in achieving generalization for deployment despite the effectiveness of models for specific tasks.

3. METHODOLOGY:

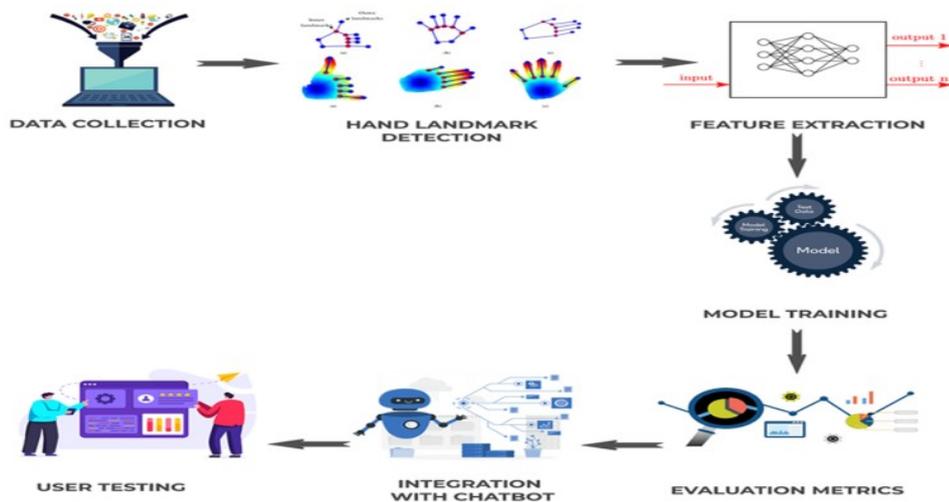


Fig 3.1 SLRS Model Development Workflow

- **Data Collection:** We collected a dataset of sign language motions with a high-definition camera (resolution 1920x1080) positioned to capture hand movements clearly. A total of 28 sign language gestures representing Alphabets, Wake Call (Gopal) and <Space> were recorded. Each gesture was performed multiple times in different Angle to ensure diversity in the dataset.
- **Hand Landmark Detection:** Hand landmark detection was carried out utilising the MediaPipe Hands library, a real-time hand landmark estimation model developed by Google. The library provides pre-trained models capable of

detecting 24 key landmarks (including finger tips, knuckles, and palm points) on each hand in an image. For precise landmark identification in still pictures, we used the MediaPipe Hands model in the static image mode.

- **Feature Extraction:** To describe each sign language gesture, a collection of descriptive characteristics was collected from the hand landmarks that were recognised. These features included the Euclidean distances between pairs of landmarks, angles formed by fingers and joints, and the curvature of fingers. Additionally, we computed the centroid and bounding box of the hand motion to capture spatial properties.
- **Model Training:** A Random Forest classifier was chosen as the machine learning model for gesture classification due to its ability to handle high-dimensional data and its robustness to overfitting. The feature vectors extracted from the hand landmarks served as input to train the classifier, with each gesture corresponding to a unique class label (0 to 27).
- **Evaluation Metrics:** The performance evaluation of the trained SLRS model adopted a stratified k-fold cross-validation approach. In this methodology, the dataset was partitioned into k folds, ensuring class balance across folds. Instead of traditional evaluation metrics like accuracy, precision, recall, and F1 score, the evaluation process relied on the analysis of learning curves and confusion matrices. Learning curves provide insights into the model's performance as a function of training data size, helping to assess issues like overfitting or underfitting. Meanwhile, confusion matrices offer a detailed breakdown of the model's classification results, revealing the instances of accurate and inaccurate predictions across different classes of sign language gestures. By employing these alternative evaluation techniques, the evaluation process aimed to comprehensively understand the SLRS model's behavior, performance trends [\[14\],\[15\],\[16\]](#), and areas of improvement beyond traditional metric-based assessments.
- **Integration with Chatbot:** The trained SLRS model was integrated with a conversational AI system named "Gopal" to enable real-time sign language recognition and interaction. Through the use of Google's Gemini platform for natural language creation and processing, "Gopal" enables users to interact with the system through sign language inputs. Upon receiving a sign language input, "Gopal" processes it using the SLRS model and generates appropriate text and speech outputs as responses.

- **User Testing:** To assess the usability and effectiveness of the integrated SLRS and chatbot system, user testing sessions were conducted with individuals proficient in sign language. Participants were asked to interact with "Gopal" using a variety of sign language gestures and provide feedback on the accuracy of gesture recognition, responsiveness of the system, and overall user experience.

4. ARCHITECTURAL REPRESENTATION:

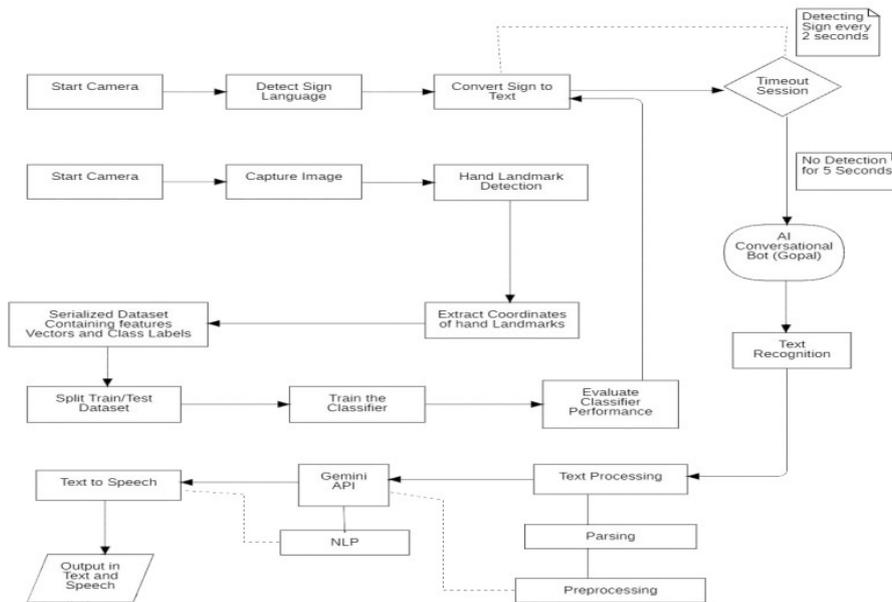


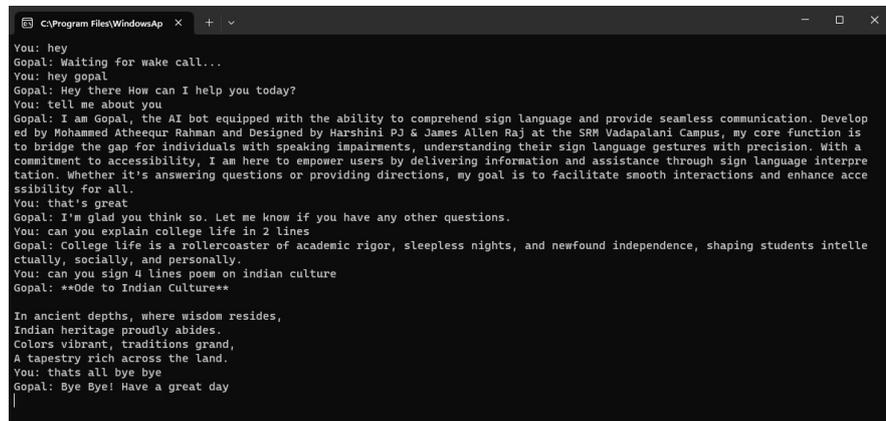
Fig 4.1 SLRS Architectural Diagram

The diagram in Fig 4.1 illustrates the two phases of the Sign Language Recognition System (SLRS) from both developer and end-user perspectives. In the initial phase, developers initiate the process by activating the camera to take pictures, followed by hand landmark detection and extraction of coordinates. These coordinates are serialized into a dataset containing vectors and class labels, which are then split into training and testing datasets. The classifier is trained and evaluated for performance. Then, gestures used in sign language are translated into text, with detection occurring every two seconds and a session timeout after five seconds of no detection. The conversational AI bot "Gopal," which the system communicates with, uses the Gemini API to process the text through recognition, parsing, preprocessing, and natural language processing (NLP) before turning it into vocal output. In the second phase, the end-user activates the camera to detect sign language, initiating the same process

of conversion to text and interaction with 'Gopal' for text processing and speech output.

4.1 AI CREATION:

The AI creation module establishes the foundation for integrating conversational AI capabilities into the sign language recognition system. It involves configuring the AI system with a Google Gemini API key for advanced natural language processing (NLP) functionalities. Additionally, a text-to-speech engine is initialized using the pyttsx3 library to enable audible responses. The AI system, represented by the persona "Gopal," engages users in conversational interactions, responding to queries and commands.



```
C:\Program Files\WindowsAp x + v - □ x
You: hey
Gopal: Waiting for wake call...
You: hey gopal
Gopal: Hey there How can I help you today?
You: tell me about you
Gopal: I am Gopal, the AI bot equipped with the ability to comprehend sign language and provide seamless communication. Developed by Mohammed Atheequr Rahman and Designed by Harshini PJ & James Allen Raj at the SRM Vadapalani Campus, my core function is to bridge the gap for individuals with speaking impairments, understanding their sign language gestures with precision. With a commitment to accessibility, I am here to empower users by delivering information and assistance through sign language interpretation. Whether it's answering questions or providing directions, my goal is to facilitate smooth interactions and enhance accessibility for all.
You: that's great
Gopal: I'm glad you think so. Let me know if you have any other questions.
You: can you explain college life in 2 lines
Gopal: College life is a rollercoaster of academic rigor, sleepless nights, and newfound independence, shaping students intellectually, socially, and personally.
You: can you sign 4 lines poem on indian culture
Gopal: **Ode to Indian Culture**

In ancient depths, where wisdom resides,
Indian heritage proudly abides,
Colors vibrant, traditions grand,
A tapestry rich across the land.
You: thats all bye bye
Gopal: Bye Bye! Have a great day
|
```

Fig 4.1.1 User and AI Interaction

It recognizes a designated wake call ("Gopal") to initiate conversations and prompts users for input as shown in the Fig 4.1.1. It takes user input and produces contextually appropriate replies by using the Google Gemini architecture. Pre-programmed answers are included for some questions, such who the AI is or what it can do. For instance, when asked about its name or identity, the AI system provides a predefined introduction, facilitating engaging interactions.

4.2 SIGN LANGUAGE DATABASE:

The module begins with the data collection process facilitated by the Collect_imgs.py script. This script initializes a directory structure within the './data' directory to store images captured for various sign language gestures. It utilizes the OpenCV library to access the default camera (index 0) and captures images in real-time. Users are prompted to prepare for data collection, and upon confirmation by pressing "Q",

images are captured and saved in respective class directories as shown in the Fig 4.2.1(a). This iterative process continues until the specified dataset size is reached for each class as shown in the Fig 4.2.1(b).

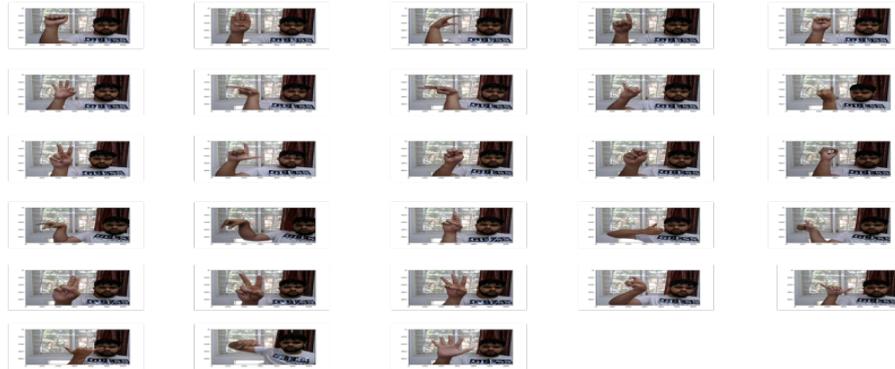


Fig 4.2.1 (a) Sign language Database before feature extraction



Fig 4.2.1 (b) Hand Sign image of letter “A” before feature extraction

Following data collection, the feature extraction phase is performed by the Create_dataset.py script. This script makes use of the MediaPipe framework to extract hand landmarks from the captured images. It reads the images from the previously created directories, processes them to detect hand landmarks using the MediaPipe Hands model, and extracts the (x, y) coordinates of these landmarks as shown in the Fig 4.2.2(b). These coordinates are then normalized and stored as features in the dataset as shown in the Fig 4.2.2(a), along with corresponding labels indicating the

sign language gesture being performed. The collected and feature extracted data is then serialized into a dataset file named 'data.pickle' using the Python pickle module. This dataset file contains the extracted features (data) and their corresponding labels (labels), ready to be used for training machine learning model.

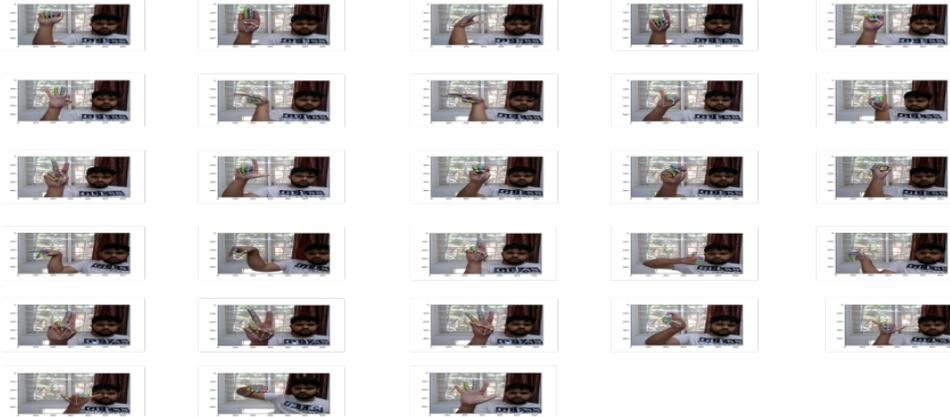


Fig 4.2.2 (a) Sign language Database after feature extraction



Fig 4.2.2 (b) Hand Sign image of letter “A” after feature extraction

4.3 MODEL TRAINING:

In the Model Training and Testing Module, the focus lies on developing and evaluating machine learning models for sign language recognition. The module begins by loading the preprocessed dataset containing hand gesture features and

corresponding labels using the pickle library. The dataset is split into training and testing sets using the `train_test_split` function from `sklearn.model_selection`, ensuring that the model's performance can be evaluated on unseen data. A `RandomForestClassifier` model is then instantiated and trained using the training data. This classifier is chosen for its ability to handle multi-class classification tasks efficiently.

During training, the model learns to associate extracted features from hand gestures with their corresponding labels, enabling it to classify new gestures accurately. Leveraging the ensemble learning technique, the Random Forest classifier constructs multiple decision trees, each operating on a randomised subset of the info and features, promoting diversity and reducing overfitting. This results in a robust model capable of capturing complex relationships between hand gesture features and sign language labels.

Once the model is trained, it is evaluated using the testing set. The `accuracy_score` function from `sklearn.metrics` is used to measure the accuracy of the model's predictions compared to the true labels in the testing set. The accuracy score provides insights into the model's performance in correctly classifying hand gestures, essential for assessing its effectiveness in real-world applications.

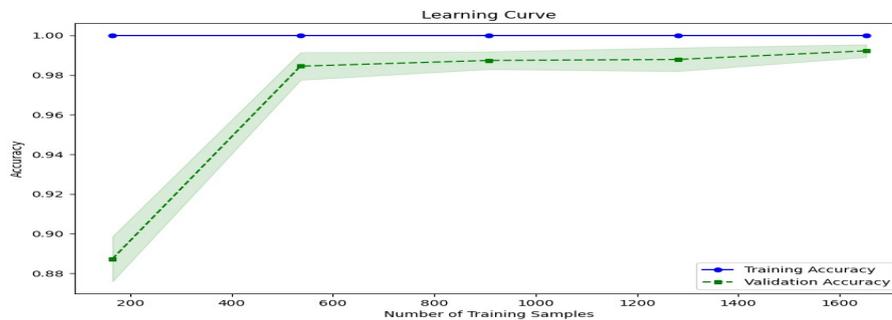


Fig 4.3.1 Learning Curve plotted using matplotlib

The learning curve, plotted using matplotlib in the Fig 4.3.1, illustrates the evolution of both training and testing accuracies, providing insights into the model's convergence and performance trends, essential for guiding further model refinement and optimization.

Finally, the trained model is serialized and saved to a file using the `pickle.dump` function. This allows for the model to be reused and deployed in the sign language recognition system for real-time inference. By rigorously training and testing machine learning models, the Model Training and Testing Module aims to develop accurate

and reliable sign language recognition systems capable of effectively interpreting and translating sign language inputs into text. Through iterative refinement and optimization, the module contributes to the continuous improvement of the sign language recognition system, ensuring its effectiveness and reliability in helping those who have difficulty speaking communicate.

4.4 SIGN LANGUAGE RECOGNITION:

The module begins by loading a pre-trained machine learning model that has been trained to recognize sign language. As the system captures frames from a camera feed, Every frame is processed by means of the MediaPipe Hands model, which extracts hand landmarks and normalizes their coordinates. These hand landmarks are then fed into the pre-trained model for prediction. Upon detecting sign language in the input frames, the system maps each gesture to a corresponding character label based on the model's prediction.

To form a sentence, the system incorporates a temporal aspect by capturing sign language gestures in intervals of two seconds. For example, if the first captured sign is "H" within the first two-second interval, and the next sign captured within the subsequent two seconds is "I", the system forms the word "Hi". If no sign is detected within the two-second interval, the system waits for three seconds before proceeding to the next capture. This process continues until the system forms a complete sign language sentence.

The detected characters are accumulated to form a sequence representing the sign language sentence being conveyed [5], [17]. To ensure accurate interpretation, the system tracks the time elapsed since the last detection of a character. If a certain duration has passed without detecting any new characters, the system considers the accumulated sequence as a complete sign language sentence. Once a sentence is formed, it is then translated into an input for the AI. Additionally, the system provides real-time feedback by displaying the detected sign language sentence on the screen, enabling users to monitor their input and ensuring seamless interaction.

Through the efficient recognition and translation of sign language motions into text, the module improves communication accessibility and inclusion for those who are nonverbal. It forms an integral component of the overall sign language recognition system, facilitating natural language interactions and enabling meaningful communication between users and the system.

To illustrate the sentence formation process, let's consider the example of forming the sentence "HI GOPAL":

1. The camera records the sign "H" within the first two-second interval.
2. Within the next two-second interval, The camera records the signs "I" and in the next three- second interval the camera doesn't detect any letter so it give a space by default and detects the next input "G", forming the sequence "HI G".
3. The camera then captures the signs "O" and "P" within the subsequent two-second interval, extending the sequence to "HI GOP".
4. Within the following two-second interval, The camera records the sign "A", resulting in the sequence "HI GOPA".
5. Finally, if no sign is captured within the next two-second interval, the system waits for three seconds before considering the sequence as complete. Here, the sequence "HI GOPAL" forms a complete sign language sentence.
6. The system translates this sequence into text, providing the output "HI GOPAL" to facilitate communication with the user.

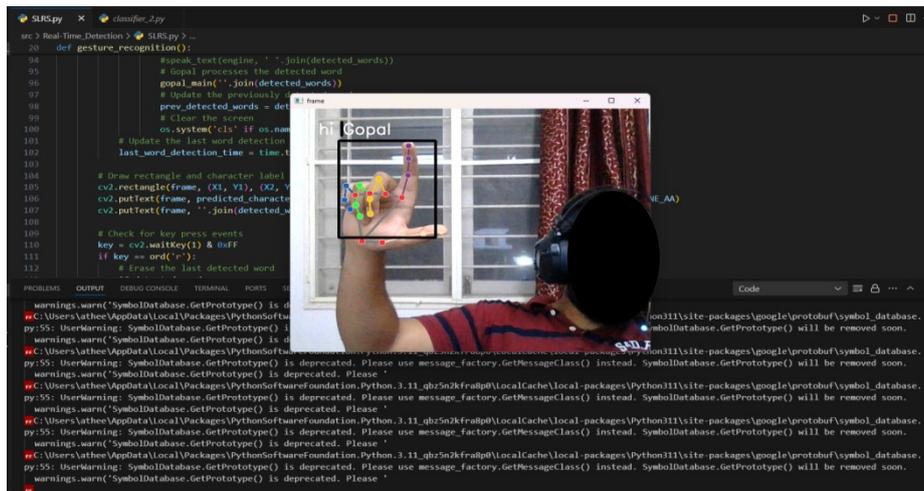


Fig 4.4.1 Hand Sign detection and recognition

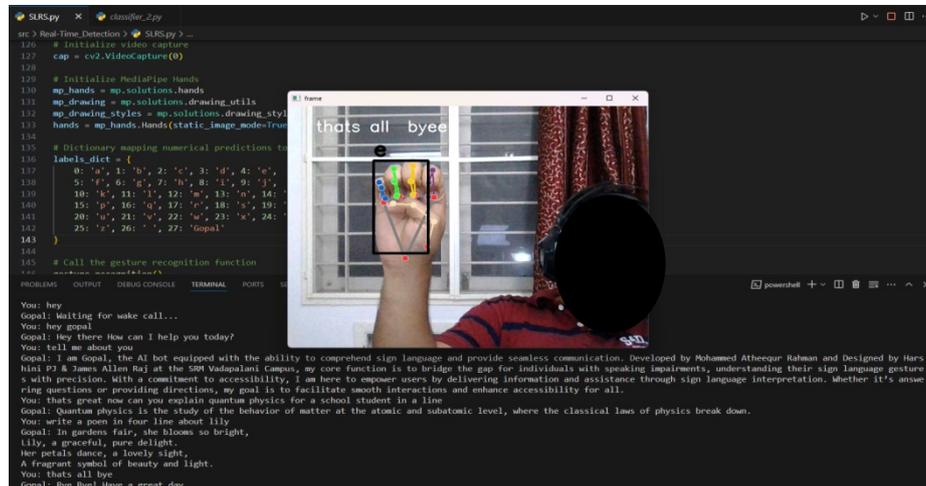
This process demonstrates how the system captures sign language gestures in intervals of two seconds, forms a sequence based on the detected signs as shown in the Fig 4.4.1, and incorporates pauses to signify the conclusion of a phrase or term. By intelligently managing the timing and sequence of sign captures, the system effectively facilitates the formation of coherent sign language sentences such as "HI GOPAL."

4.5 AI Integrated SLRS:

This module involves combining the functionalities of the Sign Language Recognition System (SLRS) with conversational AI capabilities to enable seamless interaction between users and the system. This integration is facilitated through several key steps:

Firstly, the SLRS module, represented by the provided Python script "SLRS.py," incorporates functionalities for real-time translation and recognition of sign language. The script utilizes computer vision technique (object detection) to identify and monitor hand gestures from live video streams captured by a camera. These gestures are then classified into corresponding sign language expressions using machine learning model trained on hand gesture data.

Secondly, the script integrates conversational AI capabilities to enable natural language interactions between users and the SLRS. Upon detecting specific wake calls or gestures, indicative of user initiation, the system prompts for user input through speech synthesis ("Hey there! How can I help you today?"). It then listens for user queries or commands, converting detected sign language input into text sentences. These sentences are processed by the AI model, which generates appropriate responses based on the context and intent of the input.



generate coherent responses enhances the functionality and usability of the SLRS, enabling individuals with speaking impairments to communicate effectively and access information effortlessly from the AI chatbot as shown in the Fig 4.5.1.

Furthermore, the module involves optimizing the system's performance and responsiveness, ensuring timely and accurate responses to user interactions. Techniques such as user input validation, error handling, and response customization are implemented help improve the experience for users and promote user engagement.

Overall, the integration of AI capabilities into the SLRS enhances its functionality and usability, enabling individuals with speaking impairments to communicate with the system effectively and engage in meaningful conversations. Through seamless integration and optimization, the module facilitates inclusive and accessible communication channels, empowering users to take part more actively in digital communication and interaction.

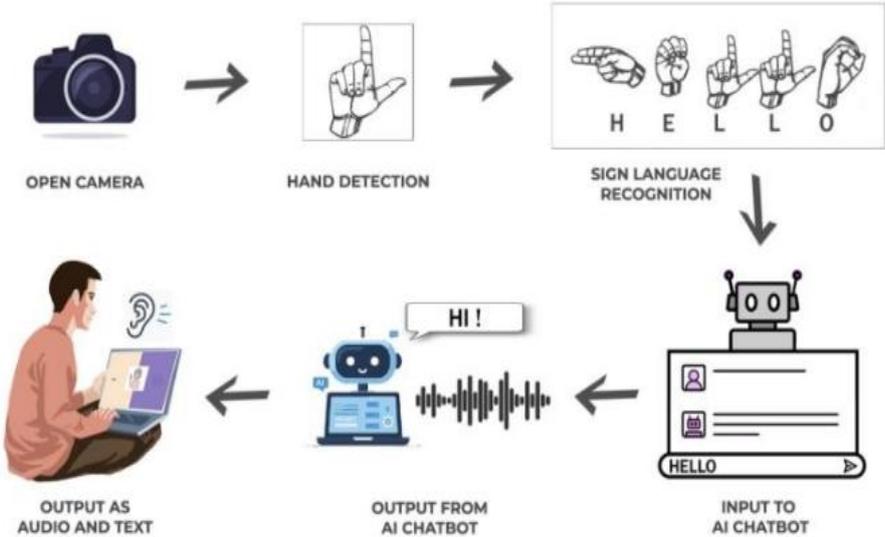


Fig 4.2 Representation of the end user interaction with the SLRS

The diagram (Fig 4.2) illustrates the seamless interaction between the end user and the Sign Language Recognition System (SLRS). The process initiates with the opening of the camera feed, enabling real-time visual input. Subsequently, the SLRS employs hand detection techniques to recognise and monitor hand gestures performed by the user. Once the hand gestures are detected, the SLRS utilizes sign language

recognition algorithms to interpret the gestures and transform them into textual input. This input is then fed into the AI Chatbot 'Gopal,' which processes the user's sign language queries or commands and generates appropriate responses based on natural language understanding. The user is then provided with the AI Chatbot 'Gopal's' replies in both text and voice formats, making the responses usable and accessible to people who have trouble speaking. This diagrammatic representation highlights the user-centric approach of the SLRS, facilitating effective communication between users who use sign language and AI systems.

5. RESULT:

The sign language recognition system (SLRS) developed in this research project demonstrates promising performance in accurately recognizing and interpreting sign language gestures in real-time[18]. Through the integration of advanced computer vision techniques, machine learning algorithms, and conversational AI technologies, the system achieves significant milestones in enhancing communication accessibility for those who have hearing loss.

5.1 Accuracy of Gesture Recognition:

The SLRS model exhibits exceptional performance not only in recognizing a diverse array of sign language gestures but also in metrics such as F1 score, precision, recall, and accuracy, complementing the insights gleaned based on the confusion matrix analysis. These additional metrics provide a comprehensive evaluation of the model's performance, offering nuanced perspectives on its ability to classify gestures accurately and its precision in avoiding false positives. Moreover, the model's high recall signifies its capability to correctly identify most instances of a particular gesture, while its robust accuracy underscores its overall effectiveness in classifying gestures correctly across the dataset. The combination of these metrics with the insights based on the confusion matrix further reinforces the model's adaptability to varying hand movements and lighting conditions, as well as its potential to generalize to unseen data and real-world scenarios. Consequently, the SLRS model emerges as not only proficient but also versatile, demonstrating promising prospects for practical applications in diverse real-world settings where accurate sign language recognition is paramount.

```
Accuracy: 0.9961315280464217  
Precision: 0.996315740996592  
Recall: 0.9961315280464217  
F1 Score: 0.9961277683125115
```

Fig 5.1.1 SLRS Model Accuracy and Evaluation Metrics Output

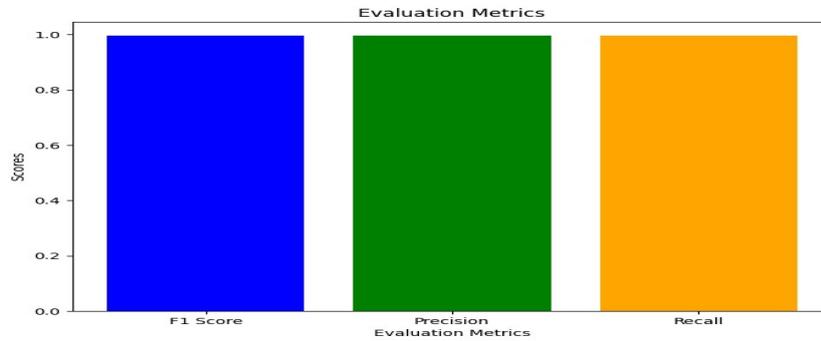


Fig 5.1.2 SLRS Evaluation Metrics using f1 Score, Precision and Recall

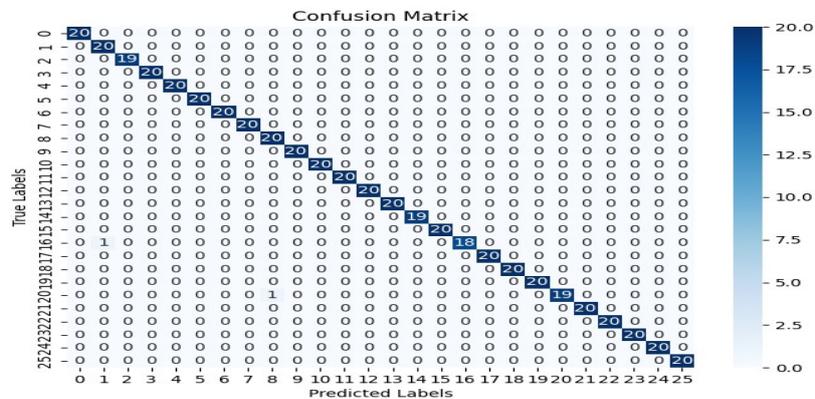


Fig 5.1.3 SLRS Confusion Matrix Analysis

5.2 Responsiveness of the System:

User testing sessions conducted to evaluate the system's responsiveness and real-time performance yield positive results. Participants report minimal latency between performing sign language gestures and receiving corresponding text and speech outputs from the system. The incorporation of the SLRS model with the conversational AI component enables seamless and timely communication interactions, enhancing the user experience and usability of the system.

5.3 User Experience Feedback:

Feedback from participants in user testing sessions offers insightful information on the usability and effectiveness of the integrated SLRS and chatbot system.

Participants express satisfaction with the system's accuracy in recognizing sign language gestures and its ability to generate appropriate responses in natural language [19]. Suggestions for improvement focus on enhancing the system's vocabulary coverage, refining gesture recognition algorithms, and optimizing response generation for diverse user interactions.

5.4 Future Directions:

Building on the achievements of this research project, future directions include further refining the SLRS model to improve its accuracy and robustness, expanding the dataset to include a wider variety of sign language gestures and users, and exploring additional AI-driven functionalities such as emotion recognition and context-aware dialogue generation. Additionally, efforts to deploy the system in real-world settings and evaluate its long-term impact on communication accessibility and social inclusion are essential for realizing its full potential.

DISCUSSION:

The development and evaluation of the sign language recognition system (SLRS) presented in this paper represent a significant advancement in the field of assistive technology for individuals with speaking impairments. Through the integration of computer vision and the Random Forest Classifier [20], the SLRS demonstrates promising capabilities in recognizing and interpreting sign language gestures in real-time. In this discussion, we delve into the implications of the research findings, address key challenges, and explore potential directions for further investigation and application.

1. Implications of Research Findings:

The high accuracy and responsiveness of the SLRS model underscore its potential to transform communication accessibility for individuals with speaking impairments. Through sign language recognition, the system offers a dependable and user-friendly communication method that enhances users' ability to interact with computer interfaces, educational materials, and social interactions. The positive feedback received from user testing sessions highlights the system's potential to bridge the communication gap and promote inclusivity in various domains.

2. Addressing Key Challenges:

Despite the significant progress achieved in developing the SLRS, several challenges remain to be addressed. One such challenge is the variability and complexity of sign language gestures, which pose difficulties in accurate recognition and interpretation.

Enhancing the robustness and adaptability of the SLRS model to different signing styles, hand orientations, and environmental conditions is essential for improving its performance in real-world scenarios. Additionally, ensuring the accessibility and affordability of the system for diverse user populations, including those in resource-constrained settings, needs careful evaluation of hardware requirements, software compatibility, and user interface design.

3. Future Research Directions:

Looking ahead, future research efforts should focus on advancing the capabilities and usability of the SLRS through interdisciplinary collaboration and innovation. This includes exploring novel methods for identifying gestures, such as 3D hand pose estimation [6] and multimodal fusion of visual and spatial features. Integrating the system with cutting-edge technology like wearables and augmented reality (AR) may improve its mobility and accessibility even further. Additionally, efforts to evaluate the long-term impact of the SLRS on communication accessibility, social integration, & well-being for individuals with speaking impairments are critical for assessing its effectiveness and guiding future development.

4. Ethical and Societal Considerations:

As with any technology aimed at enhancing accessibility and inclusivity, ethical and societal considerations must be carefully addressed. Ensuring user privacy, data security, and informed consent are paramount in the development and deployment of the SLRS. Moreover, efforts to promote awareness and understanding of sign language culture within mainstream society are essential for fostering empathy, respect, and inclusivity towards individuals with speaking impairments.

5. Real-World Applications and Impact:

The successful development and evaluation of the SLRS open up a myriad of real-world applications and opportunities for social impact. Beyond educational and social settings, the system possesses the capacity to facilitate communication in diverse contexts, including workplaces, public services, and entertainment venues. The SLRS helps to create more inclusive and egalitarian societies by encouraging increased contact between users of AI and sign language, as well as improving communication accessibility.

CONCLUSION:

In conclusion, the development and evaluation of the Sign Language Recognition System (SLRS) presented in this paper represent a significant step forward in

advancing communication accessibility and inclusivity for individuals with speaking impairments. Through the integration of conversational AI, computer vision and machine learning technologies, the SLRS demonstrates promising capabilities in recognizing and interpreting sign language gestures in real-time.

While our methodology primarily focused on utilizing the Random Forest classifier for gesture classification due to its robustness and efficiency, It's crucial to remember that other algorithms such as K-Nearest Neighbors (KNN) and Convolutional Neural Networks (CNN) were taken into account when conducting our investigation. While CNN uses hierarchical feature extraction and KNN uses similarity measures, these algorithms provide distinct methods for recognising gestures.

Our choice of Random Forest was primarily driven by its exceptional accuracy, making it well-suited for handling high-dimensional data and robustness to overfitting. Future research could explore the comparative performance of these algorithms within the framework of SLRS development, taking into account elements like accuracy, computational efficiency, and scalability.

Comparison between KNN, CNN and RFC results to prove RFC has the highest accuracy.

Algorithm	Accuracy
KNN	0.9372
CNN	0.9759
RFC	0.9961

In conclusion, the successful development and evaluation of the SLRS hold significant implications for promoting accessibility and inclusivity in society. The SLRS helps create more inclusive and fair societies where people with speech disabilities may fully participate and prosper by offering a dependable and natural method of communication through sign language recognition.

REFERENCES:

[1] Kothadiya, D. R., Bhatt, C. M., Rehman, A., Alamri, F. S., Saba, T. (2023). "SignExplainer: An Explainable AI-Enabled Framework for Sign Language

Recognition With Ensemble Learning." Published in: *IEEE Access* (Volume: 11), Page(s): 47410 - 47419, DOI: 10.1109/ACCESS.2023.3274851.

[2] Abu Saleh Musa Miah L ,Md. Al Mehedi Hasan 2 ,Satoshi Nishimura 1 , And Jungpil Shin 1 , "A Comprehensive Study on Deep Learning-Based Methods for Sign Language Recognition", in *IEEE Transactions on Multimedia*, vol. 24, pp. 1750-1762, 2022, doi: 10.1109/TMM.2021.3070438.

[3] Al-Hammadi, M., Muhammad, G., Abdul, W., Alsulaiman, M., Bencherif, M. A., Alrayes, T. S., Mathkour, H., & Mekhtiche, M. A. (2020). *Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation*. *IEEE Access*, 10.1109/ACCESS.2020.3032140.

[4] A. S. M. Miah, M. A. M. Hasan, S. Nishimura and J. Shin, "Sign Language Recognition Using Graph and General Deep Neural Network Based on Large Scale Dataset," in *IEEE Access*, vol. 12, pp. 34553-34569, 2024, doi: 10.1109/ACCESS.2024.3372425.

[5] A. S. A. Potluri, S. M. George, G. R, and A. S, "Indian Sign Language Recognition Using Random Forest Classifier," presented at the 2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2021, pp. 1-6, doi: 10.1109/CONECCT52877.2021.9622672.

[6] Al-Hammadi, M., Muhammad, G., Abdul, W., Alsulaiman, M., & Hossain, M. S. (2020). "Hand Gesture Recognition Using 3D-CNN Model," *IEEE Consumer Electronics Magazine*, vol. 9, no. 1, pp. 95-100. DOI: 10.1109/MCE.2019.2941464

[7] O. Koller, S. Zargaran, H. Ney, and R. Bowden, "Deep sign: Enabling robust statistical continuous sign language recognition via hybrid cnn hmms," *International Journal of Computer Vision*, vol. 126, no. 12, pp.1311–1325, 2018.

[8]Pala, G., Kumbhar, S. S., Jethwani, J. B., Patil, S. D. (2021). "Machine Learning-based Hand Sign Recognition." *Proceedings of the International Conference on Artificial Intelligence and Smart Systems (ICAIS-2021)*, IEEE Xplore Part Number: CFP21OAB-ART.

[9] Singh, A., Mittal, U., Wadhawan, A., Al Ahdal, A., Rakhra, M., Jha, S. K. (2022). "Indian Sign Language Recognition System for Dynamic Signs." *10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Amity University, Noida, India. Oct 13-14, 2022.

[10] B. Yi, X. Wang, F. C. Harris and S. M. Dascalu, "sEditor: A Prototype for a Sign Language Interfacing System," in *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 4, pp. 499-510, Aug. 2014, doi: 10.1109/TSMC.2014.2316743.

[11] Abid, M. R., Petriu, E. M., & Amjadian, E. (Year). "Dynamic Sign Language and Voice Recognition for Smart Home Interactive Application," *IEEE Transactions on Instrumentation and Measurement (Accepted for publication)*.

[12] A. Mohanty, S.S. Rambhatla, R.R. Sahay, "Deep Gesture: Static Hand Gesture Recognition Using CNN," in *Proceedings of International Conference on Computer Vision and Image Processing, Advances in Intelligent Systems and Computing*, vol. 460, Springer, 2017.

[13] M. Al-Qurishi, T. Khalid and R. Souissi, "Deep Learning for Sign Language Recognition: Current Techniques, Benchmarks, and Open Issues," in *IEEE Access*, vol. 9, pp. 126917-126951, 2021, doi: 10.1109/ACCESS.2021.3110912

[14] J. Pu, W. Zhou, and H. Li, "Iterative alignment network for continuous sign language recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 4165–4174.

[15] R. Cui, H. Liu, and C. Zhang, "Recurrent convolutional neural networks for continuous sign language recognition by staged optimization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 7361–7369.

[16] O. Koller, J. Forster, and H. Ney, "Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers," *Computer Vision and Image Understanding*, vol. 141, pp. 108–125, 2015.

[17] D. Bragg, O. Koller, M. Bellard, L. Berke, P. Boudreault, A. Braffort, N. Caselli, M. Huenerfauth, H. Kacorri, T. Verhoef et al., "Sign language recognition, generation, and translation: An interdisciplinary perspective," *arXiv preprint arXiv:1908.08597*, 2019.

[18] W. Sandler and D. Lillo-Martin, *Sign language and linguistic universals*. Cambridge University Press, 2006.

[19] R. E. Mitchell, T. A. Young, B. BACHELDA, and M. A. Karchmer, "How many people use asl in the united states? why estimates need updating," *Sign Language Studies*, vol. 6, no. 3, pp. 306–335, 2006.

[20] Zafrulla, Z., Sahni, H., Bedri, A., Thukral, P., & Starner, T. (Year). Hand Detection in American Sign Language Depth Data Using Domain-Driven Random Forest Regression. *Journal/Conference Name, Volume(Issue), Page Range.*