Website for seamless translation of sign language using KNN and TensorFlow

MINI PROJECT REPORT

Submitted by

Subhikshaa S **210701264**

Tamil Priya V **210701282**

in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI

ANNA UNIVERSITY: CHENNAI 600 025

MAY 2024

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI

BONAFIDE CERTIFICATE

Certified that this Report titled "**Website for seamless translation of sign language using KNN and TensorFlow**" is the bonafide work of **"Subhikshaa (210701264) and Tamil Priya V (210701282)"** who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

Karthick. V
Associate Professor,
Department of Computer Science and Engineering,
Rajalakshmi Engineering College,
Chennai – 602015

Submitted to Mini Project Viva-Voce Examination held on ___________________

Internal Examiner **External Examiner**

ABSTRACT

The Sign Language Translator is a user-friendly website designed to train the sign language that acts as a bridge for the communication between users. The core functionality lies in capturing and training individual signs with their corresponding meanings. It proceeds with a start and stop gesture to maintain a flow of end and beginning of the translation. It uses a minimum of thirty samples to get trained with more accuracy in recognizing the sign language shown. The start and stop gesture is collected and trained. The next step includes the signs that need to be trained with their meaning with atleast thirty samples each. This translator depicts how a machine learning model works by training the samples and testing on the data given by the user. It detects the sign by using the trained samples and gives the result as text with computer generated speech in the background. One can sign multiple words with one gesture and copy the translated text with a click of a button. Its purpose is to allow users to communicate more effectively with their computers and other people. The use of KNN (K-Nearest Neighbours) algorithm contributes to gesture recognition of sign language when trained with samples. Users can initiate and conclude the translation process through start and stop gestures trained by them thus giving a controlled flow. Previous research work in this area has models that can translate trained symbols in 2.3 to 3.1 seconds with a confidence level of 84%. We targeted to focus on this time interval, confidence level and also showcase the project with a straight-forward user interface. It also serves more as an educational tool in the view of communication between deaf people and the normal people providing a way to vanish the indifferences. This translator additionally provides the feature of copying the text translated from recognizing the sign language which offers functionalities such as sharing it with others or integrating it to a document. The auditory reinforcement of the translated text is another notable feature of this translator offering a clear vocal representation.

ACKNOWLEDGEMENT

Initially we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavour to put forth this report. Our sincere thanks to our Chairman **Mr. S.MEGANATHAN, B.E, F.I.E.**, our Vice Chairman **Mr. ABHAY SHANKAR MEGANATHAN, B.E., M.S.,** and our respected Chairperson **Dr. (Mrs.) THANGAM MEGANATHAN**, **Ph.D.,** for providing us with the requisite infrastructure and sincere endeavoring in educating us in their premier institution.

Our sincere thanks to **Dr. S.N. MURUGESAN, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete our work in time. We express our sincere thanks to **Dr. P. KUMAR, Ph.D.**, Professor and Head of the Department of Computer Science and Engineering for his guidance and encouragement throughout the project work. We convey our sincere and deepest gratitude to our internal guide, **KARTHICK V** Professor, Department of Computer Science and Engineering. Rajalakshmi Engineering College for his valuable guidance throughout the course of the project.

SUBHIKSHAA S - 210701264

TAMIL PRIYA V - 210701282

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

KNN	K-Nearest Neighbors
LSTM	Long Short-Term Memory
Open CV	Open-Source Computer Vision Library
DDR4	Double Data Rate 4
GPU	Graphics Processing Unit

CHAPTER 1
INTRODUCTION

1.1 GENERAL

The Sign Language Translator facilitates communication between deaf and hearing individuals by training and recognizing gestures using the K-Nearest Neighbors algorithm. It provides text and speech outputs, allows multiple-word gestures, and includes features for copying translated text. This user-friendly tool enhances accessibility, learning, and inclusivity.

1.2 OBJECTIVE

The Sign Language Translator website facilitates communication by training and recognizing sign language gestures. Using the K-Nearest Neighbors algorithm, it captures and interprets signs with a minimum of thirty samples for accuracy. Users initiate and end translations with start and stop gestures, producing text and computer-generated speech outputs. This tool aids communication between deaf and hearing individuals, provides educational value, and includes features for copying and sharing translated text, enhancing accessibility and integration into documents.

1.3 EXISTING SYSTEM

Existing system includes sign language translation using OpenCV, Deep learning algorithmic techniques and libraries to train the model with a specific set of datasets or symbols and uses it to test separate samples. Several researchers also employ perception neuron modal to translate sign languages of various communication languages. Vision based systems have also been proposed for image distinction. Image processing has to be very accurate in translating them to words. However, there are also significant challenges posed over this.

1.4 PROPOSED SYSTEM

Our proposed solution aims to provide a flexible method or platform where users of any category can train the model with customized signs and can assign values to their signs. The use of adobe illustrator files ensures high quality designs are being rendered. They process the input symbols received from the user and transform them into suitable formats for training. The system uses KNN algorithm and TensorFlow libraries to detect and translate the signs. As our solution is employed as a website which can be made available anywhere in the world and can be used anytime. It proves to be a reliable solution with high user desirability and at the same time ensures feasibility.

CHAPTER 2
LITERATURE SURVEY

The work by Prashant and Khushboo [7] demonstrates the recognition of various ISL hand gestures and alphabets. They have performed detailed analysis of the human-computer interaction approaches and other image processing techniques and have identified some key difficulties including costs and budgets. Their system uses OpenCV and python in training the model. They tried to provide an answer for almost all the underlying questions they have identified. Utilizing color techniques, they have tried to provide a stable solution.

Srivastava, Amisha Gangway, Richa Mishra, Sudhakar Singh [8] in their research on sign language recognisers, have concentrated more in the recent advancements and have confined their research to the latest possible date of publishing. They have considered various inclusion and exclusion parameters for filtering the studies. Their inclusion criterias were based on the language of the paper, the algorithm used with the paper and the keywords and their exclusion criterias were based on the papers that did not implement deep learning models and also the ones that were not fully accessible.

Siming He [9] in his research work has studied the sign language applications and have developed a hand locating, sign language recognition system based on popular and familiar signs using neural networks. Using LSTM coding and decoding, he has established a sign language recognition framework which handles and processes the data. They aim to increase the recognition accuracy. This paper uses a hand locating network to overcome the problem of RGB color recognition.

Yang sang, Mengru Liu and Feilu Wang [10] have employed the use of pressure sensors and have integrated 7 different categories of gestures, each category or group having a set of gesture symbols together in a wrist band. Using the response time of the sensors, they have developed their solution.

Achenbach, Laux and Muller [2] has demonstrated the use of hand gloves in recognising hand symbols. They have employed multiple classifiers that can filter out samples and noncontributing features to train an efficient model.

Pathan, Biswas and Yasmin [11] have used a multi-headed convolutional network in recognising known words through hand gestures. They have used a variety of datasets in training the model for basic symbols and sentences.

Gonzalez [12] has concentrated his research on the area of processing images digitally for training the model in predicting results of hand movements. They claim that proper training using digital image processing techniques can render fruitful results.

Cozza, Fiasca and Martini [1] have gone through the mechanism behind hearing loss and how people with hearing difficulty process our oral communication. Their study about the working underlying the deaf people have provided a different perspective in this area of research.

Ewe, Lee and others [13] have attempted to limit the overfitting problems on sign language recognition of US datasets using Random Forest algorithms. They have also analyzed the research papers published before them and have mentioned their complexities.

Kakizaki, Miah, Hirooka and Shin [14] have developed a solution for the community of deaf people in japan using MediaPipe. They have used RGB cameras for carefully collecting the dataset. They have mentioned extending this work as a global project in their future work.

Ngugen and Wang [15] have highlighted the use of latitude and longitudinal axes for location detection. They have predicted the speech using the angles in facial movement and have proven a significant accuracy level. They claim that smaller changes in facial angles can also be detected and can prove helpful in the overall recognition process.

CHAPTER 3

SYSTEM DESIGN

3.1 DEVELOPMENT ENVIRONMENT

3.1.1 HARDWARE SPECIFICATIONS

This project uses minimal hardware but in order to run the project efficiently without any lack of user experience, the following specifications are recommended

Table 3.1.1 Hardware Specifications

Processor	Intel Core i5
RAM	4GB or above (DDR4 RAM)
GPU	Intel Integrated Graphics
Hard Disk	6GB
Web camera	Minimum 1080p
Audio Speakers	Minimum 5W

3.1.2 SOFTWARE SPECIFICATIONS

The software specifications in order to execute the project has been listed down in the below table. The requirements in terms of the software that needs to be pre-installed and the languages needed to develop the project has been listed out below.

Table 3.1.2 Software Specifications

FRONT END	HTML, CSS, JavaScript, Adobe Illustrator files
BACK END	Node JS
FRAMEWORKS	Tensor Flow
CODE EDITOR	Visual Studio

3.2 SYSTEM DESIGN

3.2.1 ARCHITECTURE DIAGRAM

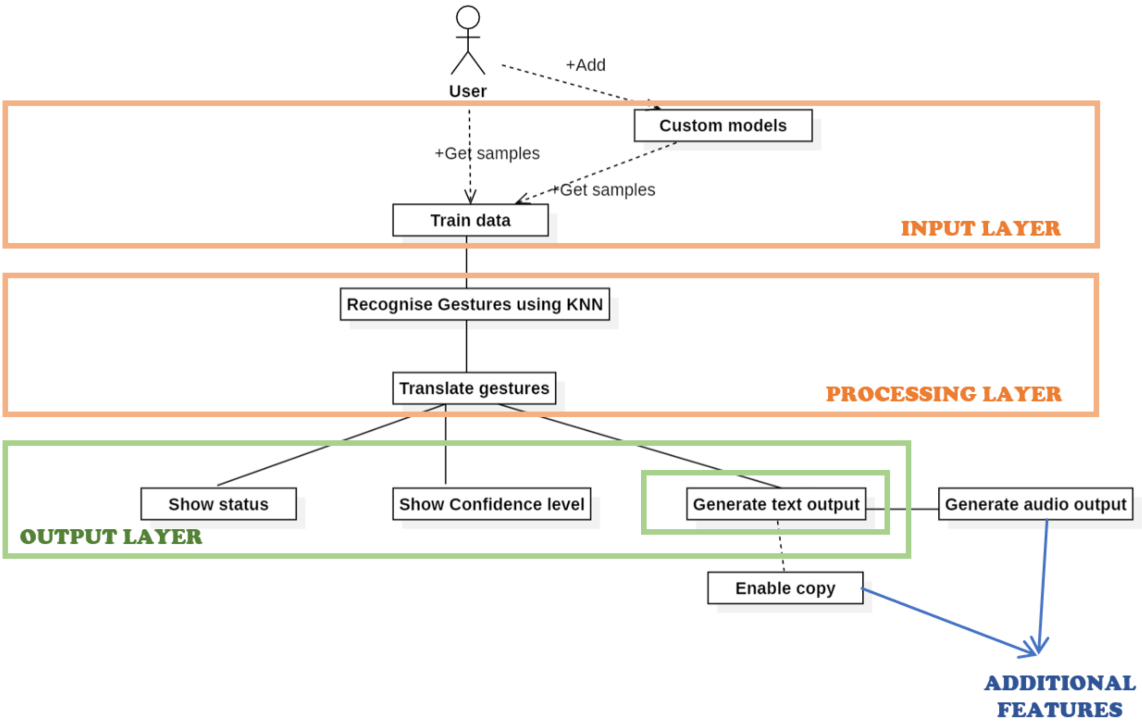

Fig 3.2.1 Architecture Diagram

PRE-PROCESSING:

Before the Sign Language Translator application can effectively translate sign language gestures into text, it undergoes several preprocessing steps to prepare the input data for training and prediction. The preprocessing primarily focuses on optimizing the webcam feed and organizing the training data. Firstly, the application initializes the webcam and configures its settings to ensure a stable and consistent video input stream. This step involves adjusting parameters such as resolution, frame rate, and camera angle to optimize gesture recognition accuracy. Additionally, the application performs background subtraction or segmentation to isolate the user's hand gestures from the surrounding environment, reducing noise and improving the model's robustness to varying lighting conditions.

Furthermore, the preprocessing phase involves organizing the training data into distinct classes corresponding to different sign language gestures. This step entails labeling and annotating the training examples to associate each gesture with its corresponding class label. The training data may also undergo data augmentation techniques, such as flipping, rotating, or scaling the images, to increase the diversity of training examples and enhance the model's generalization capabilities. Additionally, the preprocessing phase may involve normalizing the input data to ensure consistency in feature scales and improve model convergence during training.

TRAINING SET:

The training set for the Sign Language Translator comprises labeled examples of sign language gestures captured through the webcam feed. Each gesture is associated with a specific class label, such as "Start" or "Stop," allowing the model to learn the correspondence between hand movements and linguistic meaning. The training set is carefully curated to include diverse examples of each gesture, ensuring the model's ability to accurately recognize and translate sign language gestures into text.

CHAPTER 4
PROJECT DESCRIPTION

4.1 MODULE DESCRIPTION

A. Data collection:

The initial step of the website is to collect sample data that is getting images of signs from the user to train the start and stop gesture. It requires a minimum of 30 samples to train each of the start and stop gestures. These pre required signs are collected to proceed to the next step to train other signs. The start and stop signs serve as markers to indicate beginning and end of a sentence to determine the prediction of sign language as a conversation.

B. Adding custom words:

The model can be trained with samples for their own words or sentences and can view their added words with their trained gestures below the training area. All these new words or sentences require a minimum of 30 samples as mentioned earlier. All these samples can be continuously captured or any faulty image taken in between can be clear according to a last in first out fashion.

C. Translation initiation:

Once done with adding samples for their corresponding gesture, users can click on the start translation button to initiate the translation process. They have to start with the start gesture trained in the previous step to get started with the process.

D. Translation process:

During the translation process, each predicted word with the confidence level appears straight before the video location. For models trained in a well light environment and moderately lit environment, confidence level will be much closer to 100%. In most of the cases, it achieves a confidence level of 99%.

E. Audio generation:

While the translation is being seamlessly done, an additional voice that converts the text to speech also gets triggered. It reads aloud the content being translated after the start symbol in a second’s time.

F. Live status availability:

Whether translation is being done or not, there will be a status message pop up those shows what is happening with the model. The users can use this as a quick hold for getting through what is happening at the second with the model. It shows if it is translating or not at the left corner of the screen.

G. Translation completion:

Showing the stop symbol to the translator marks the end of the translation process. Anytime during this or the previous stage, users can go back and add new words and train gestures by providing live samples. Until the user returns back to the homepage, all the trained data is retained in and to other navigable pages.

CHAPTER 5

IMPLEMENTATION AND RESULTS

5.1 IMPLEMENTATION

The codebase for the Sign Language Translator project encompasses a comprehensive implementation of a web-based application designed to facilitate real-time translation of sign language gestures into text. It leverages HTML, CSS, and JavaScript to create an intuitive user interface that guides users through the process of gesture training, prediction, and even video calls. The code is well-structured, with clear separation of concerns between different components such as the training interface, translation functionality, and video call integration. By encapsulating each feature within its respective module, the codebase promotes modularity and maintainability, allowing for easier troubleshooting and future enhancements.

Gesture Training and Prediction:

A significant portion of the code is dedicated to gesture training and prediction functionalities. It utilizes the K-nearest neighbours (KNN) algorithm to classify gestures based on input from the webcam feed. Users can train custom gestures, such as the "Start" and "Stop" gestures, by providing examples through the training interface. The code dynamically updates the training count and displays visual feedback to users, such as checkmarks indicating successful training. Once trained, the system predicts gestures in real-time and translates them into text, providing immediate feedback to users. Additionally, the code handles edge cases, such as pausing and resuming the prediction process, to ensure a seamless user experience.

User Interface and Interaction:

The code prioritizes user experience by implementing an intuitive interface and enabling smooth interaction with the application. It features responsive design elements to accommodate various screen sizes and devices, ensuring accessibility for all users. The user interface guides users through each stage of the translation process, from training gestures to engaging in real-time conversations. Clear status indicators and instructional prompts help users understand the current stage of the application and provide guidance on next steps. Moreover, the code includes interactive components, such as buttons for initiating translation or switching between training and prediction modes, enhancing user engagement and usability.

5.1.1 RESULT:

Thus, our website was able to provide a swifter response than the other systems that we referred to. Also, as the website has an intuitive User interface that can be accessed by people of any age, the target audience for our project has been expanded to a greater extent.

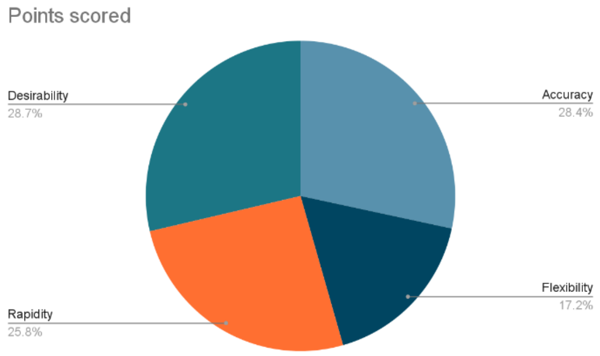

Fig 5.1.1 Features that were observed in our solution

The metrics that we focused on for transition which was the time taken to recognize and translate the gesture was reduced and we achieved a greater accuracy with a quicker and yet stable response time. The previous works that we considered mostly were constrained to specific categories of people being the audience or target users. Also, their model worked for a common vocabulary dataset and does not allow for any customization options.

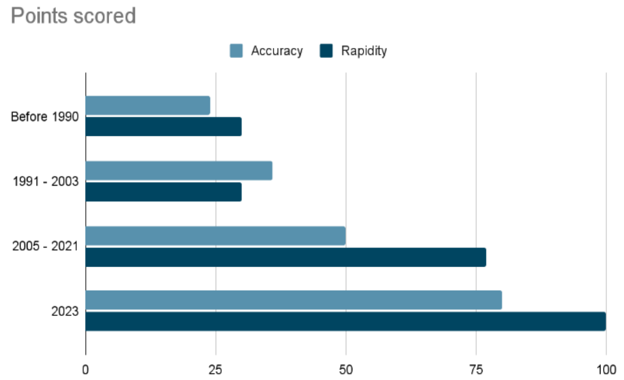

Fig 5.1.2 Comparison of previous research and their results (Accuracy and Rapidity in prediction) over the years

5.2 OUTPUT SCREENSHOTS

In our experimentation, we observed noteworthy outcomes, particularly in terms of response time. We witnessed a remarkably swift reaction, with processing occurring within the range of 5 to 6 microseconds. This performance far surpasses that of the models we had previously examined, showcasing a significant advancement in efficiency and speed. The substantial reduction in processing time signifies a notable improvement in overall system performance, promising enhanced productivity and effectiveness in real-world applications.

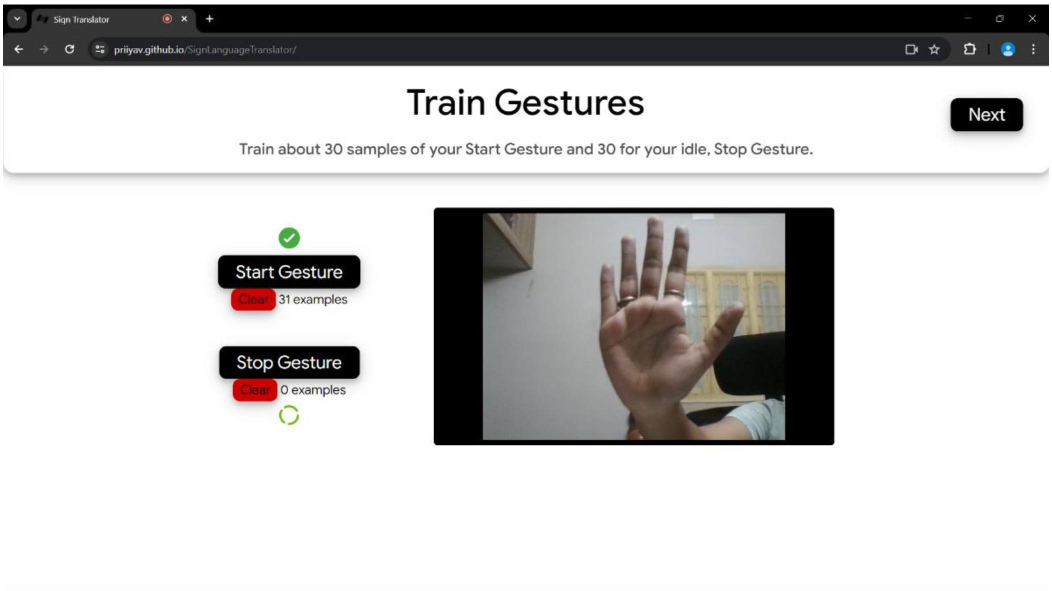

Fig 5.2.1 Train Gestures

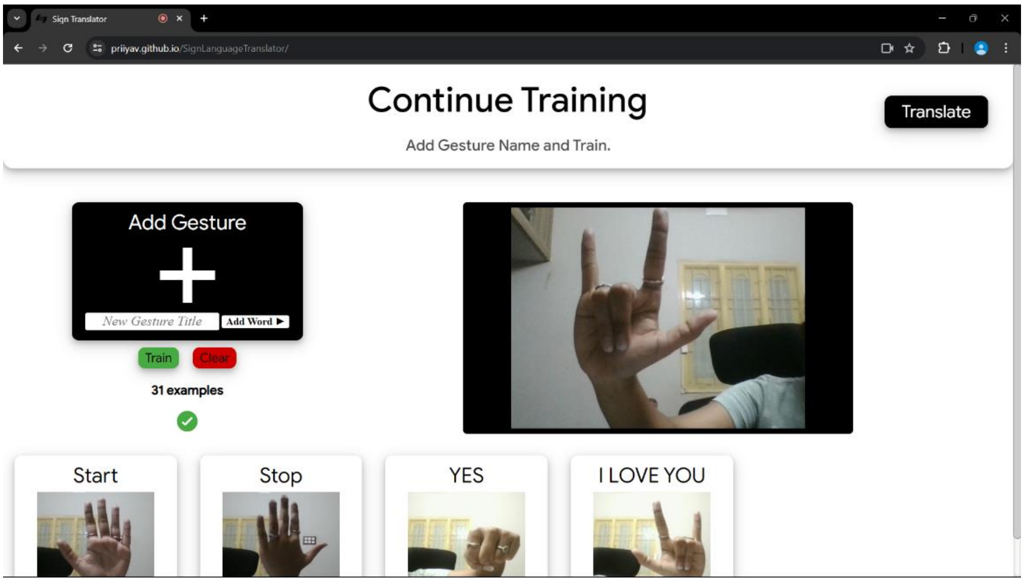

Fig 5.2.2 Status of the model

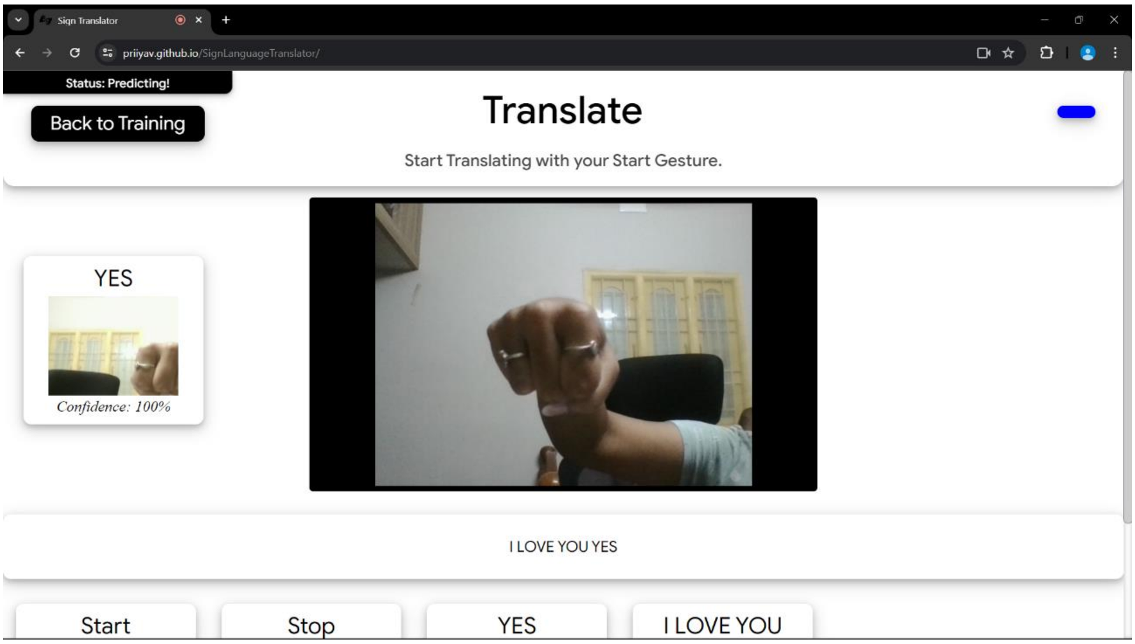

Fig 5.2.3 Confidence Level

CHAPTER 6
CONCLUSION AND FUTURE ENHANCEMENTS

6.1 CONCLUSION

Thus, our solution aims to provide a reliable and easy-to-use platform accessible to everyone, regardless of their hearing abilities. The Sign Language Translator leverages advanced gesture recognition and speech synthesis technologies to ensure seamless communication between deaf and hearing individuals. Its user-friendly interface facilitates the effortless transfer of recognized text, making information sharing simple and efficient.

The core functionality of our translator lies in its ability to accurately capture and train individual signs with their corresponding meanings. By requiring a minimum of thirty samples for each sign, we ensure high accuracy in recognition. Users can initiate and conclude translations with start and stop gestures, maintaining a controlled flow during the communication process. This adaptability allows for the translator to be customized according to the user's specific needs, providing a personalized experience. Ultimately, the Sign Language Translator serves as a vital resource for fostering understanding and eliminating communication barriers between deaf and hearing individuals.

6.2 FUTURE ENHANCEMENTS

Future enhancements for the Sign Language Translator could include the integration of real-time translation for live conversations, expanding the database to support multiple sign languages, and incorporating machine learning algorithms for improved accuracy and adaptability. Additionally, developing a mobile application version would increase accessibility and convenience. Enhanced user customization features, such as personalized sign libraries and voice options, could further tailor the experience. Collaborations with educational institutions and organizations for the deaf could provide valuable feedback and ensure the tool meets diverse needs, ultimately broadening its impact and effectiveness in fostering inclusive communication.

REFERENCES

[1] A. Cozza, V. M. Di Pasquale Fiasca, and A. Martini, “Congenital Deafness and Deaf-Mutism: A Historical Perspective,” Children, vol. 11, no. 1, Dec. 2023, doi: 10.3390/children11010051.

[2] P. Achenbach, S. Laux, D. Purdack, P. N. Müller, and S. Göbel, “Give Me a Sign: Using Data Gloves for Static Hand-Shape Recognition,” Sensors , vol. 23, no. 24, Dec. 2023, doi: 10.3390/s23249847.

[3] J. Pataky, E. C. Demalis, J. Shelly, K. Miller, Z. M. Moore, and M. E. Vidt, “Use of a factor analysis to assess biomechanical factors of American Sign Language in native and non-native signers,” J. Biomech., vol. 165, p. 112011, Mar. 2024.

[4] S. Escalera, I. Guyon, and V. Athitsos, Gesture Recognition. Springer, 2017.

[5] A. Camurri and G. Volpe, Gesture-Based Communication in Human-Computer Interaction: 5th International Gesture Workshop, GW 2003, Genova, Italy, April 15-17, 2003, Selected Revised Papers. Springer, 2011.

[6] E. Bergeron, R. Valdez, C. J. Moreland, R. Wang, T. Knight, and P. Kushalnagar, “Community Health Navigators for Cancer Screening Among Deaf, Deafblind, and Hard of Hearing Adults Who Use American Sign Language,” J. Cancer Educ., Feb. 2024, doi: 10.1007/s13187-024-02416-x.

[7] Verma, Prashant, and Khushboo Badli. "Real-Time Sign Language Detection using TensorFlow, OpenCV, and Python." International Journal for Research in Applied Science & Engineering Technology (IJRASET) 10.V (2022): May2022.

[8] ““Sign Language Recognition System using TensorFlow Object Detection API.” Accessed: May 06, 2024. [Online]. Available: https://arxiv.org/pdf/2201.01486

[9] S. He, “Research of a sign language translation system based on deep learning,” in 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), IEEE, Oct. 2019. doi: 10.1109/aiam48774.2019.00083.

[10] Y. Song, M. Liu, F. Wang, J. Zhu, A. Hu, and N. Sun, “Gesture Recognition Based on a Convolutional Neural Network-Bidirectional Long Short-Term Memory Network for a Wearable Wrist Sensor with Multi-Walled Carbon Nanotube/Cotton Fabric Material,” Micromachines (Basel), vol. 15, no. 2, Jan. 2024, doi: 10.3390/mi15020185.

[11] R. K. Pathan, M. Biswas, S. Yasmin, M. U. Khandaker, M. Salman, and A. A. F. Youssef, “Sign language recognition using the fusion of image and hand landmarks through multi-headed convolutional neural network,” Sci. Rep., vol. 13, no. 1, p. 16975, Oct. 2023.

[12] Gonzalez, Digital Image Processing,2/e. Pearson Education India, 2002.

[13] E. L. R. Ewe, C. P. Lee, K. M. Lim, L. C. Kwek, and A. Alqahtani, "LAVRF: Sign language recognition via Lightweight Attentive VGG16 with Random Forest," PLoS One, vol. 19, no. 4, p. e0298699, Apr. 2024.

[14] M. Kakizaki, A. S. M. Miah, K. Hirooka, and J. Shin, "Dynamic Japanese Sign Language Recognition Throw Hand Pose Estimation Using Effective Feature Extraction and Classification Approach," Sensors , vol. 24, no. 3, Jan. 2024, doi: 10.3390/s24030826.

[15] J. Nguyen and Y. C. Wang, "A Classification Model Utilizing Facial Landmark Tracking to Determine Sentence Types for American Sign Language Recognition," Conf. Proc. IEEE Eng. Med. Biol. Soc., vol. 2023, pp. 1–4, Jul. 2023.

[16] M. Perea-Trigo, C. Botella-López, M. Á. Martínez-Del-Amor, J. A. Álvarez-García, L. M. Soria-Morillo, and J. J. Vegas-Olmos, "Synthetic Corpus Generation for Deep Learning-Based Translation of Spanish Sign Language," Sensors , vol. 24, no. 5, Feb. 2024, doi: 10.3390/s24051472.

[17] Y. Gu, H. Oku, and M. Todoh, "American Sign Language Recognition and Translation Using Perception Neuron Wearable Inertial Motion Capture System," Sensors , vol. 24, no. 2, Jan. 2024, doi: 10.3390/s24020453.

[18] F. Amber, Vision Based Sign Language Identification System Using Facet Analysis. GRIN Verlag, 2014.

[19] L. Kane, B. K. Dewangan, and T. Choudhury, Challenges and Applications for Hand Gesture Recognition. IGI Global, 2022.

[20] M. Gil-Martín, M. Villa-Monedero, A. Pomirski, D. Sáez-Trigueros, and R. San-Segundo, "Sign Language Motion Generation from Sign Characteristics," Sensors , vol. 23, no. 23, Nov. 2023, doi: 10.3390/s23239365.

[21] H. Alsolai, L. Alsolai, F. N. Al-Wesabi, M. Othman, M. Rizwanullah, and A. A. Abdelmageed, "Automated sign language detection and classification using reptile search algorithm with hybrid deep learning," Heliyon, vol. 10, no. 1, p. e23252, Jan. 2024.

[22] A. Ben Haj Amor, O. El Ghoul, and M. Jemni, "Sign Language Recognition Using the Electromyographic Signal: A Systematic Literature Review," Sensors , vol. 23, no. 19, Oct. 2023, doi: 10.3390/s23198343.

[23] Å. Elwér and J. Andin, "Geometry in the brain optimized for sign language - A unique role of the anterior superior parietal lobule in deaf signers," Brain Lang., vol. 253, p. 105416, May 2024.

[24] Y. Verma and R. S. Anand, "Gesture generation by the robotic hand for aiding speech and hard of hearing persons based on indian sign language," Heliyon, vol. 10, no. 9, p. e29678, May 2024