Website for seamless translation of sign language using KNN and TensorFlow

Team Members:

- Tamil Priya V - 210701282
- Subhikshaa S - 210701264

Table of contents

OUTCOME

COMPARISON

Website for seamless translation of sign language using KNN and TensorFlow

PROBLEM

METHODOLOGY

02 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 at least 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. The use of KNN (K-Nearest Neighbours) algorithm contributes to gesture recognition of sign language when trained with samples. The auditory reinforcement of the translated text is another notable feature of this translator offering a clear vocal representation.

03 Introduction

Sign language is crucial for those with hearing impairments, offering a means of communication through visual gestures. Interpreters bridge the gap between sign language and oral communication, but their numbers are limited. To address this, a website is proposed to automate translation, making communication accessible to all. The site boasts a user-friendly interface, simplifying navigation for a diverse user base. It aims to translate gestures into text, facilitating communication for those who rely on sign language. For individuals unable to speak, the website offers instant text conversion, enabling interaction with others and bridging communication barriers. Users can train the system with their own gestures and phrases, customizing the process to their needs. Additionally, the website aids those unfamiliar with sign language, allowing them to understand and empathize with individuals facing hearing or vocal challenges. It fosters community connections among the deaf and hard of hearing, enhancing communication and relationships. Moreover, a computerised voice system provides audio assistance and verifies translations

04

Literature Review

The work by Prashant and Khushboo [1] 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 Gangwar, Richa Mishra, Sudhakar Singh [2] in their research on sign language recognisers, have concentrated more in the recent advancements and have confined their research to the possible latest date of publishing. They have considered various inclusion and exclusion parameters for filtering the studies which included the language of the paper, the algorithm used with the paper, the keywords and excluded the papers that did not implement deep learning models and also the ones that were not fully accessible.

Siming He [3] 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 framework which handles and processes the data and aims to increase the recognition accuracy. It uses a hand locating network to overcome the problem of RGB color recognition.

Yang sang, Mengru Liu and Feilu Wang [4] 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. Using the response time of the sensors, they have developed their solution.

Achenbach, Laux and Muller [5] has demonstrated the use of hand gloves in recognising hand symbols. They have employed multiple classifiers that can filter out samples and non contributing features to train an efficient model.

Pathan, Biswas and Yasmin [6] 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 [7] 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 with digital image processing can render fruitful results.

Cozza, Fiasca and Martini [8] 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 [9] 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 [10] 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 [11] 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.

05

Research Gaps

Our research identified a critical gap in existing solutions, notably the absence of a universally accessible and user-friendly interface coupled with customizable training features and a paramount emphasis on rapid translation capabilities. Existing platforms often lacked the simplicity and intuitiveness akin to popular websites, which are renowned for their straightforward yet potent user interfaces. This gap in the market propelled our initiative to develop a solution that addresses these deficiencies, offering users a seamless and efficient experience with streamlined training customization and swift translation capabilities.

06 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. Training the model for 75 rounds and combining it with stratified k-fold cross validation, confidence level of prediction can be significantly increased. 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. Users can directly start to train the model with a start and stop symbol. For each symbol the model requires a minimum of 30 samples and more than 30 is also accepted. While the process of translation begins the user should start with the start symbol ensuring the start of the translation process and can show different gestures. If it is already trained for the shown gestures, it will be generated below the camera space and also simultaneously a audio will be played deciphering or reading the generated text aloud for the users to hear and also thereby confirm.

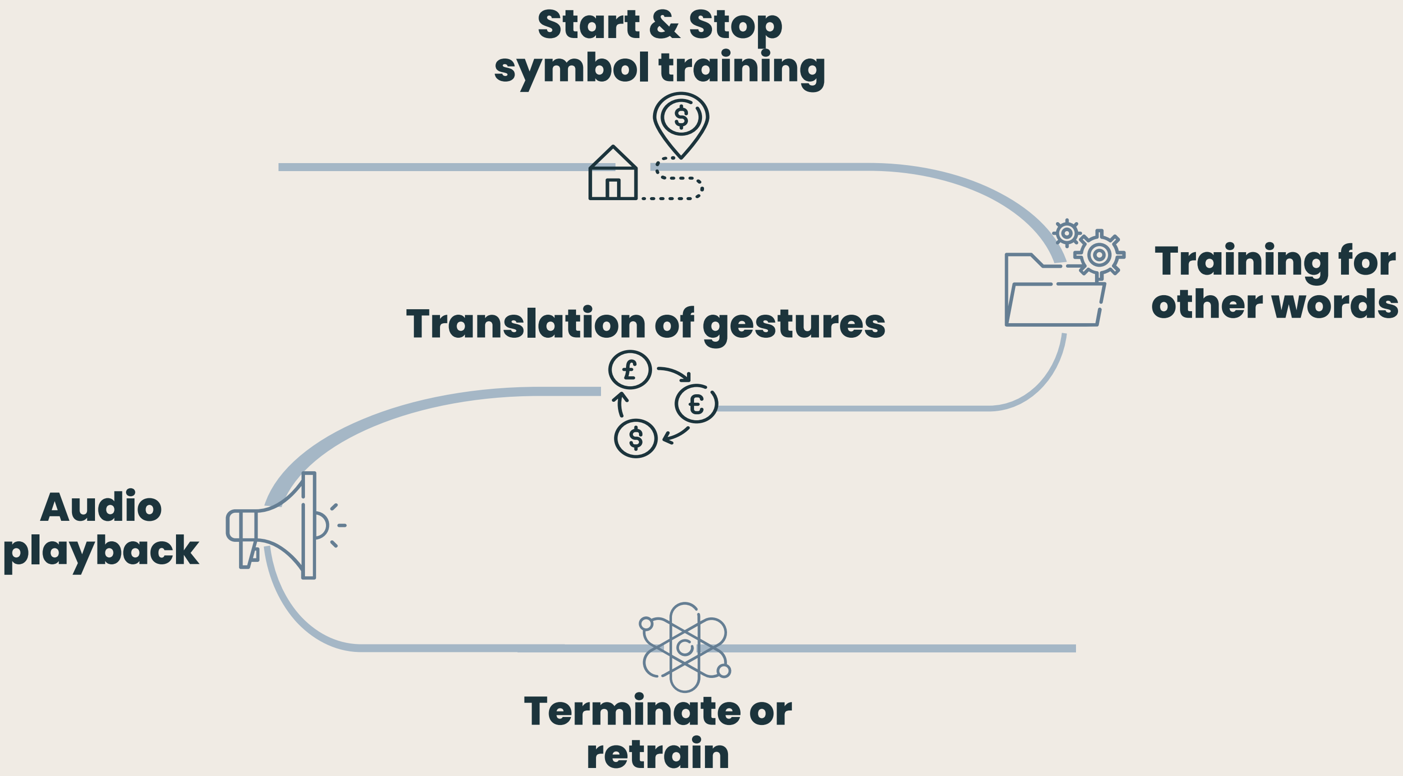

Fig. Overall process flow

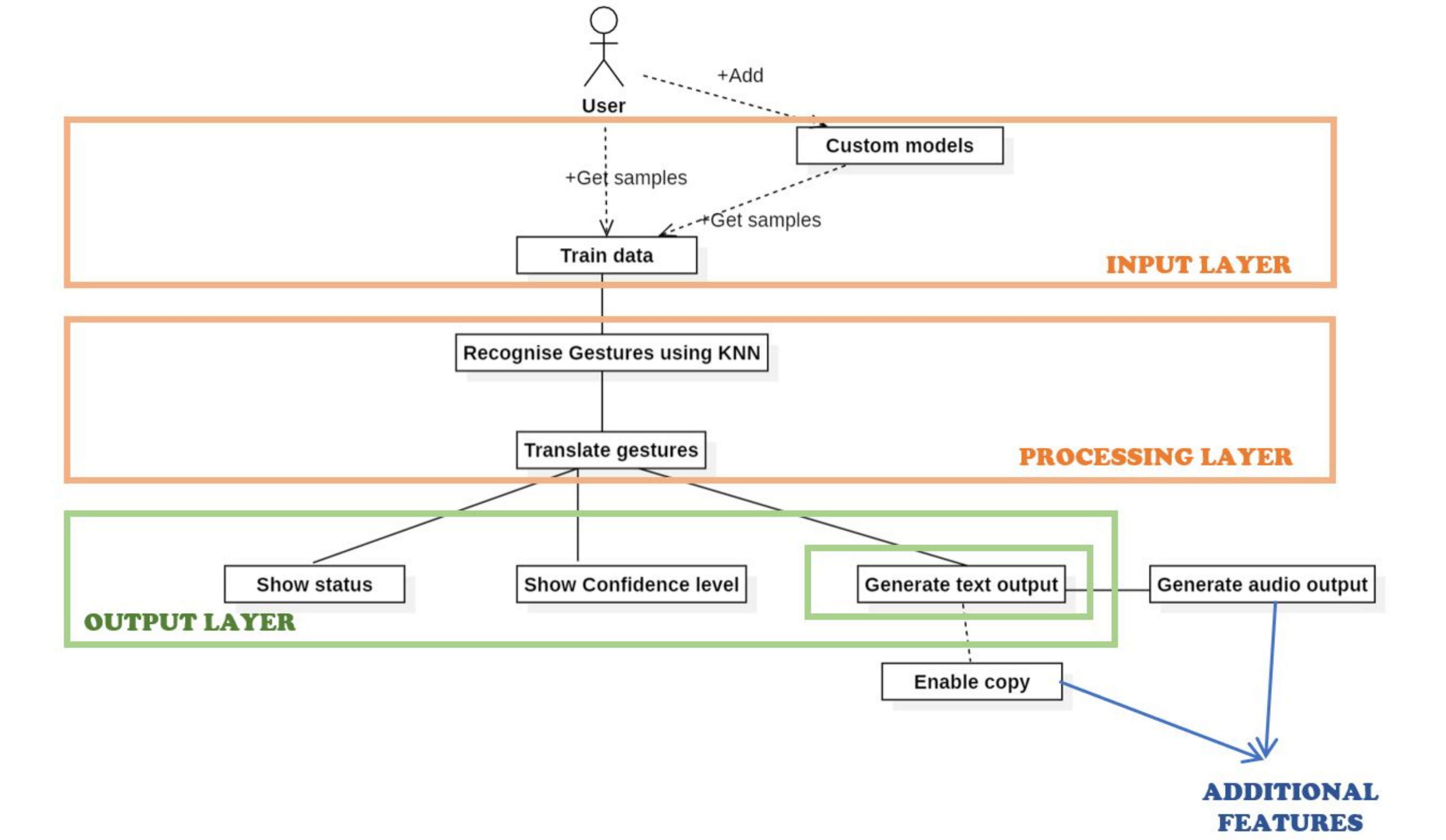

Fig. Architecture Diagram

07

Results & Discussion

Results & Discussion

Our website exhibited a significantly faster response compared to the systems we benchmarked against. The intuitive user interface made our project accessible to individuals of all ages, thereby expanding the target audience. By focusing on key transition metrics such as gesture recognition and translation time, we achieved higher accuracy rates coupled with quicker and stable response times.

In contrast to previous works, which often catered to specific user categories, our platform offers a more inclusive solution tailored to individuals with hearing difficulties. Unlike existing models limited to common vocabulary datasets, our platform provides customization options. Furthermore, we incorporated features like copy-and-paste sharing and an auxiliary audio system to enhance overall efficiency and user satisfaction. These additions contribute to a seamless user experience and facilitate easy access for all users.

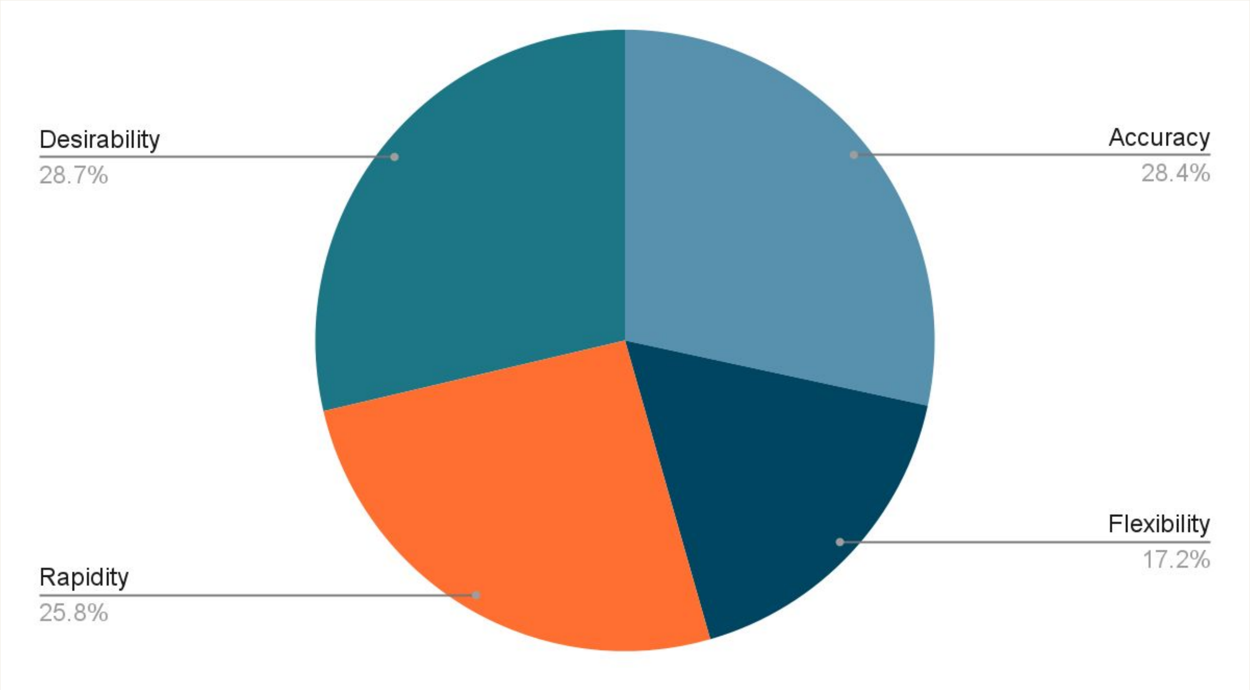

Fig. Features that were observed in our solution

08

Comparative Analysis

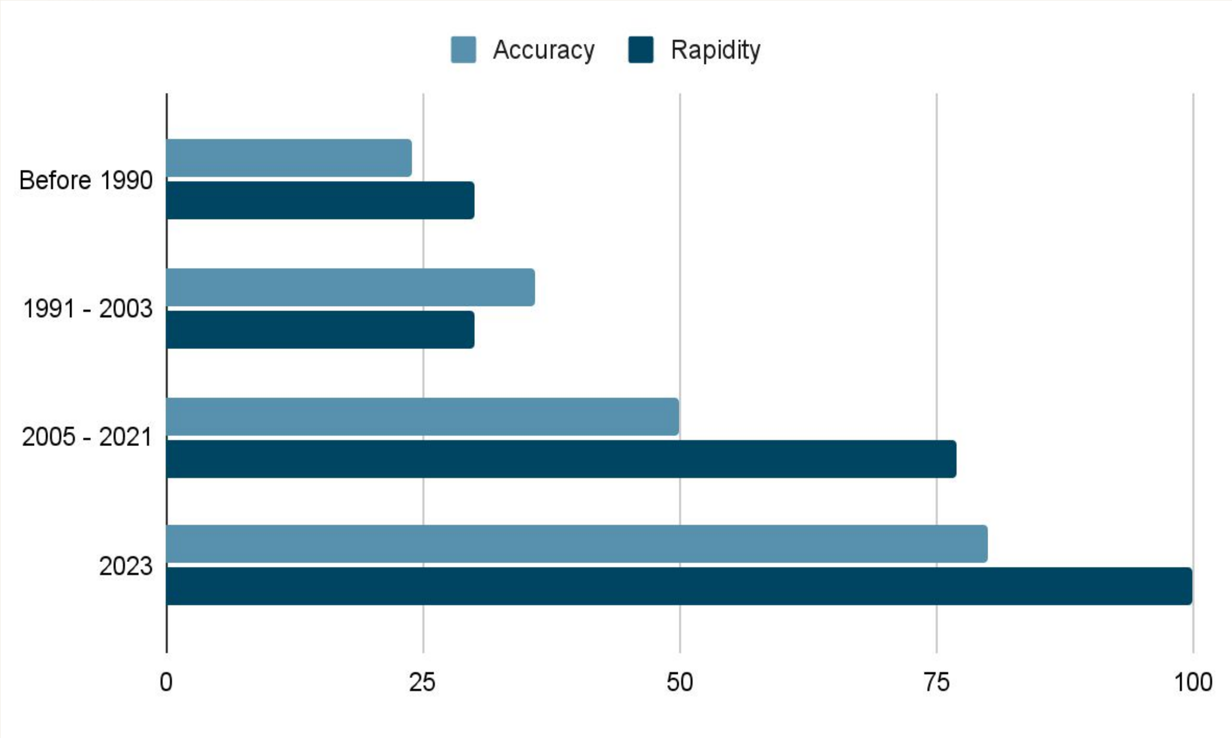

Fig. Comparison of previous research and their results (Accuracy and Rapidity in prediction) over the years

09

Conclusion & Future work

Conclusion & Future work

Our solution endeavors to furnish a dependable and user-friendly platform that caters to individuals regardless of any hearing impairments they may have. By implementing features that facilitate the seamless transfer of recognized text, our solution simplifies the sharing of information, promoting inclusivity and accessibility. Our overarching objective is to furnish an efficient medium or translator that enables rapid translation tailored to the unique needs of each user. Remarkably, our preliminary findings demonstrate a rapid response time ranging between 5 to 6 microseconds, surpassing the performance of prior models we evaluated.

As part of our future endeavors, we envision augmenting our solution with additional functionalities such as storing and saving capabilities. By integrating these enhancements, our solution will become even more invaluable to users in their daily interactions, further enhancing its utility and practicality.

10 References

[1] 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.

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

[3] 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.

[4] 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.

[5] 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.

[6] 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.

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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

Thank you!