iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 11 **Issue:** VIII **Month of publication:** Aug 2023

**DOI:** https://doi.org/10.22214/ijraset.2023.55466

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue VIII Aug 2023- Available at www.ijraset.com*

# Bi-Directional American Sign Language and English Translator Using Machine Learning and Computer Vision

Ashwin Reddy
*Department of CSE, M S Ramaiah Institute of Technology, Bangalore*

***Abstract:*** ***American sign language is a language that is used by those individuals who are physically impaired and cannot communicate through common spoken languages. Due to lack of awareness and a small population of individuals using sign language, those who rely on it as their only means of communication require human sign language translators in order to be able to communicate with the more fortunate abled individuals. However, in reality, it is not alway practically feasible to have a human translator present at all times. That is why automation of a human sign language translator's work can prove to be a more suitable alternative in a broader spectrum of use cases. A portable and practical sign language translator proves to be more socially and technically feasible than a human sign language translator.***

***Keeping these set of circumstances in mind we are developing an application that allows two-way communication between a physically impaired individual and an abled individual. The application is designed for two categories of users : those who use ASL as their primary language for communication (disabled) and those who use English as their primary language for communication (abled). The primal aim is to curb the language barrier that exists between the 2 categories of individuals in the hopes of providing better opportunities in society for the less fortunate through the use of our application. We propose to use a machine learning, computer vision based model that is capable of learning and performing the translations between English text and American Sign Language and vice versa.***

***Keywords: American Sign Language, Machine learning, Deep learning, Machine translation, Translator, Media Pipe, English***

## I. INTRODUCTION

Hearing and speech impaired people find it difficult to interact with abled people since using sign language isn't a viable option as it's not a universal language. This may lead to miscommunication between them. Currently, human sign language translators help in cases of communication between the abled and disabled people. However, having them around at all times isn't always possible. This derives the need to automate the work of human sign language translators. This is where our application proves to be beneficial and finds its use.

I developed an application that is capable of performing translation of American Sign Language to English text and back from English to American Sign Language. Our application is aimed to provide 2 way translation between the 2 types of languages. Through the use of a webcam that recognized ASL gestures and Symbols, I developed a training model that will facilitate the translation process. The application is targeted towards 2 sets of users : those individuals who use ASL as their primary language for communication and those individuals who use English as their primary language.

This project would help in facilitating simple and efficient communication between the hearing and speech impaired individuals and those who are more fortunate, while eliminating the need of a human sign language translator. This will allow and provide a sense of better social acceptance, reach and opportunities for such individuals in the society.

## II. LITERATURE SURVEY

Here I have conducted a literature survey and a market analysis of most other significant sign language translation applications available on the internet. This survey was conducted to identify missing features of currently existing sign language translators and how they can be improved. Through the market analysis we identify novelty features in our proposed application while trying to improve on currently existing applications.

Sign language translation (SLT) using frame stream density compression (FSDC) introduces an FSDC algorithm for detecting and reducing redundant similar frames, which effectively shortens the long sign sentences without losing information.

©IJRASET: All Rights are Reserved 

Then, it replaces the traditional encoder in a neural machine translation (NMT) module with an improved architecture, which incorporates a temporal convolution (T-Conv) unit and a dynamic hierarchical bidirectional GRU (DH-BiGRU) unit sequentially. The component takes the temporal tokenization information into consideration to extract deeper information with reasonable resource consumption [1].

Study of Sign Language Translation using Gesture Recognition has two modes of operation: Teach and Learn. The project uses a webcam to recognize the hand positions and sign made using contour recognition and outputs the Sign Language in PC onto the gesture made. This will convert the gesture captured via webcam into audio output which will make normal people understand what exactly is being conveyed. The project basically acquires images using the inbuilt camera of the device and performs vision analysis, functions in the operating system and provides speech outputs through the inbuilt audio device [2].

Speech To Sign Language Translator For Hearing Impaired : is a project that aims to design an application that converts the speech and text input into a sequence of sign language visuals. Speech recognition is used to convert the input audio to text and it is further translated into sign language. The proposed system is designed to overcome the troubles faced by the Indian deaf people. This system is designed to translate each word that is received as input into sign language. This project translates the words based on Indian Sign Language [3].

A paper based on Research of a Sign Language Translation System Based on Deep Learning studies hand locating and sign language recognition of common sign language based on neural networks, and the main research contents include : hand locating network based on R-CNN to recognize the sign language video or the part of the hand in the picture, and the result of recognition is passed over for processing. A 3D CNN feature extraction network and a sign-language recognition framework based on long, short time memory coding as well as decoding network are constructed for the sign language images of sequences [4].

Sign language interpreter based model paper has developed the concept of a project which enables a computer to recognize the signs used in American Sign Language (ASL) and convert the interpreted data to text for further use. Gesture recognition is done using camera capture (either through an inbuilt webcam or an externally connected device).The processing is done using Human Computer Interaction (HCI). Human computer interaction is the discipline that mainly focuses on the interaction between people and computers. This model is an attempt to be developed however provides a plethora of valuable points to take back from [5] .

Wearable device for sign language recognition : This project describes the development of a sign language translator that converts sign language into speech and text by using a wearable device. The glove-based device is able to read the movements of a single arm and five fingers. The device consists of five flex sensors to detect finger bending, and an accelerometer to detect arm motions. Based on the combination of these sensors, the device is able to identify any particular gestures that correspond to words and phrases in American Sign Language (ASL), and translate it into speech via a speaker and text which is displayed on an LCD screen [6].

Wecapable.com has developed a website based translator that converts English to ASL in a one way conversion manner. Wecapable allows for alphabet based conversion but excludes the conversion of standard commonly used phrases. Since other models and methods provide a two-way conversion from ASL to English and English to ASL they prove to be better suited for implementation in a real time conversation between an abled and disabled person [7].

Asl-dictionary.com is an application that is designed in such a way that it has predefined phrases in English and their corresponding ASL converted gestures. The application is also capable of converting English to ASL on entering text of the user's choice. However, it does not convert gestures in American Sign Language to English text for 2 way communication. In this proposed model, there is a low degree of machine learning and artificial intelligence application involved to provide efficiency and adaptability [8].

arxiv.org provides a report which mentions the use of a methodology that incorporates image processing algorithms to translate and process hand sign images of ASL, this in turn is more time consuming and resource intensive as compared to application of Computer vision for the same purpose which proves to be more efficient (and is a high-level understanding from the input digital videos with the purpose of automating human visual system). Computer Vision in theory makes a translation application faster and capable of translating multiple hand signs at once [9].

In the paper, the proposed sign language translation (SLT) includes three basic modules: object detection, feature extraction and classification and is solved using three powerful models- SSD, Inception 3 and SVM. These algorithms are applied for object detection, feature extraction and classification purposes. Firstly, Single Shot MultiBox Detection (SSD) architecture is utilized for hand detection, then a deep learning structure based on the Inception v3 plus Support Vector Machine (SVM) that combines feature extraction and classification stages is proposed to constructively translate the detected hand gestures. The criteria presented for the system were robustness, accuracy, and high speed [10].

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

## III. SYSTEM ARCHITECTURE

The System Design seeks to promote communication and collaboration among the project team members, stakeholders, and other interested parties by giving a comprehensive overview of the system's architecture. The development of software requirements, design specifications, and other system artifacts will be guided by this design, which will operate as a road map for the system's implementation and testing.

### *A. Architecture Design*

*1)* Our project follows a phase by phase approach for design and implementation. Since the problem and requirements are well known prior to the implementation of the project, following a chain of phases during implementation makes it easier to identify areas of improvement and strong or weak links. Our project's work flow follows a chain of command since it is used to convert two different languages used in a communication process.

*2)* Input to our application is obtained through the use of external or inbuilt sensors (web cameras) which is then further processed and translated to provide the corresponding matched output. Regardless of which way the communication and translation is taking place the output is always displayed on the predefined actuator (monitor).

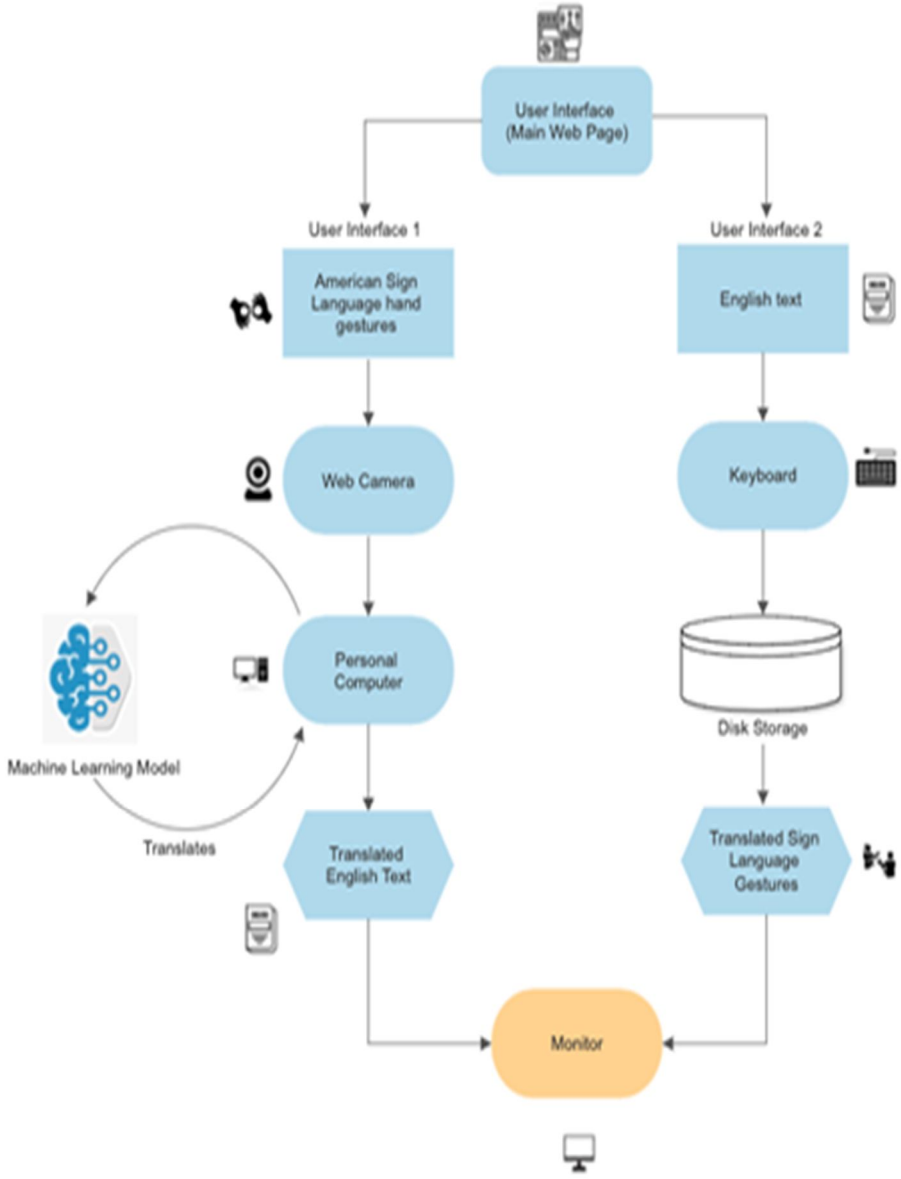



Fig 1. Architecture Diagram for the Translation model

*B. Use Case Diagram*

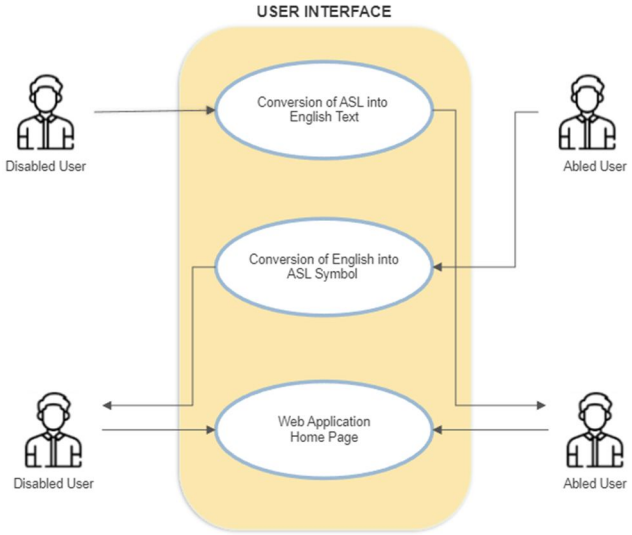



Fig 2. Use Case Diagram for the Translator model

## IV. IMPLEMENTATION

*A. Graphical User Interface*

Figure 3, shows the User Interface Mockup designs for the client-side web application. For our application's front end we are using HTML 5, CSS and PyScript to create an easy to navigate and understandable user interface through which users may browse the functionalities provided by our translator application. The User interface will provide two separate pages as a front end UI aimed at satisfying the two targeted categories of users of our application. One such page will be designed for the physically impaired set of users who wish to translate sign language into English text (the first direction of communication provided by our application) in order to communicate with an abled individual. The second page provided will be for those abled users who wish to reciprocate or communicate with a disabled individual by translating English text into American Sign Language Symbols (the second direction of communication provided by our application).

Fig 3. Web User Interface Design for the Translator

*B. Data Flow Diagram*

The low-level design of the system is depicted in Flowchart that has been designed below. It portrays the phase by phase procedure in which our project executes at run time. The Personal Computer or laptop, when given an American Sign Language hand sign input via the webcam, will attempt to recognize it by testing the recognized features on our trained machine learning model. Upon identification, the recognized alphabet or phrase is displayed as an output on the monitor in the form of English text.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

If the hand symbol or gesture is unrecognized by the application then the program continues to attempt recognition of hand symbols until and unless terminated by the user. Similarly, for converting English text to American Sign Language symbols, the input is taken from a user through the keyboard and is converted to the corresponding ASL hand symbols and displayed via the same monitor.

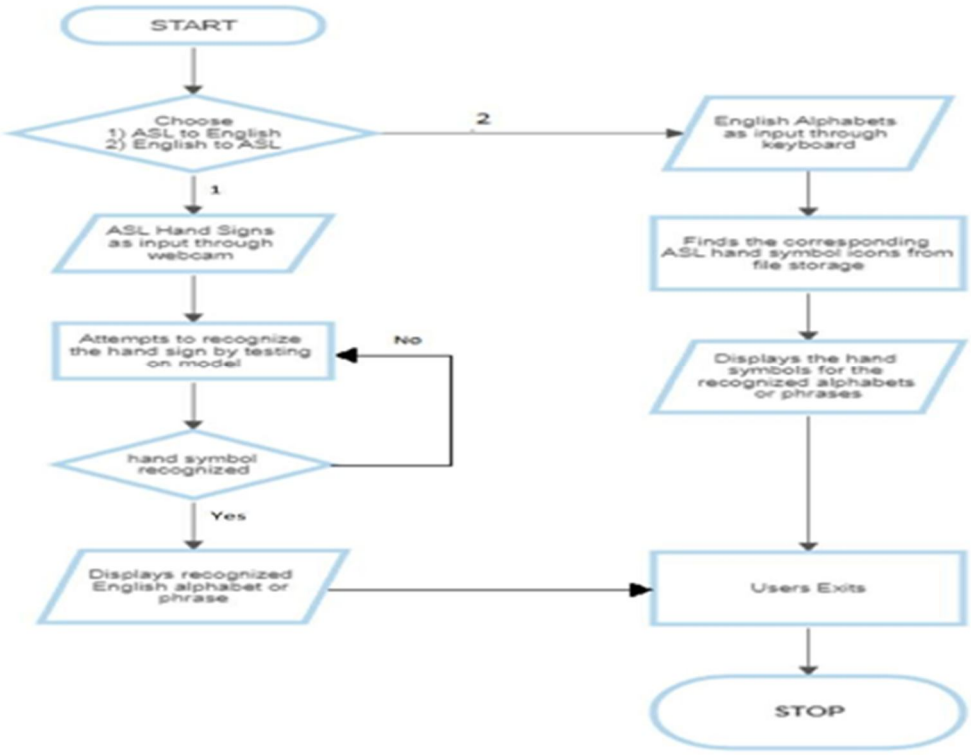



Fig 4. Process Flow for the Translation Model

*C. User Interface*

We are utilizing HTML 5, CSS, and PyScript for the frontend of our application, with the intention of crafting a user-friendly and comprehensible interface. This interface will facilitate user interaction with the various functionalities offered by our translation application. The user interface will consist of two distinct pages, each tailored to cater to specific user groups. One of these pages will be dedicated to users with physical impairments who seek to convert sign language into English text (representing the primary communication direction of our application). This enables them to effectively communicate with individuals who possess full abilities. The second page will serve users who are not physically impaired and intend to engage in communication with disabled individuals. It allows them to translate English text into American Sign Language Symbols.

*D. Hardware and Software Interfaces*

1) *Visual Studio Code*: Visual Studio Code is a source code editor made by Microsoft for Windows, Linux and MacOS. The majority of popular programming languages are supported in Visual Studio Code on a basic level. This fundamental support consists of configurable snippets, code folding, bracket matching, and syntax highlighting.
2) *Media Pipe*: The MediaPipe Framework is used to create machine learning pipelines for processing time-series data, including audio, video, and other types. This cross-platform Framework works in Desktop/Server, Android, iOS, and embedded devices. MediaPipe Toolkit comprises the Framework and the Solutions.
3) *Tensor Flow*: TensorFlow is an end-to-end open source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML powered applications.
4) *Programming Languages*: The Machine Learning model is built in Python, while the UI is built in JavaScript, HTML, and CSS.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

## V. RESULTS

### *A. ASL to English*

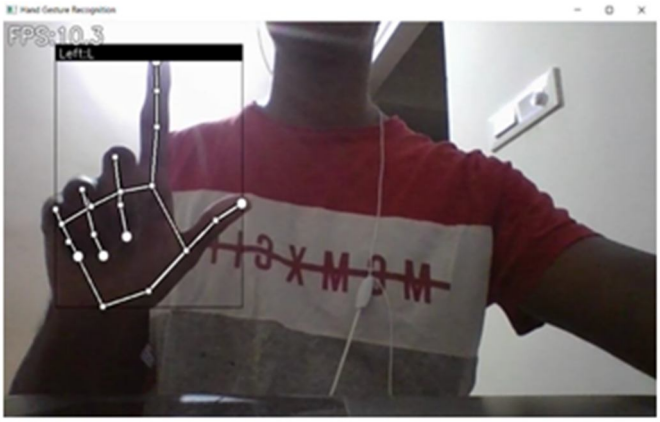



Fig 5 : ASL to English feature recognizing the alphabet 'L'

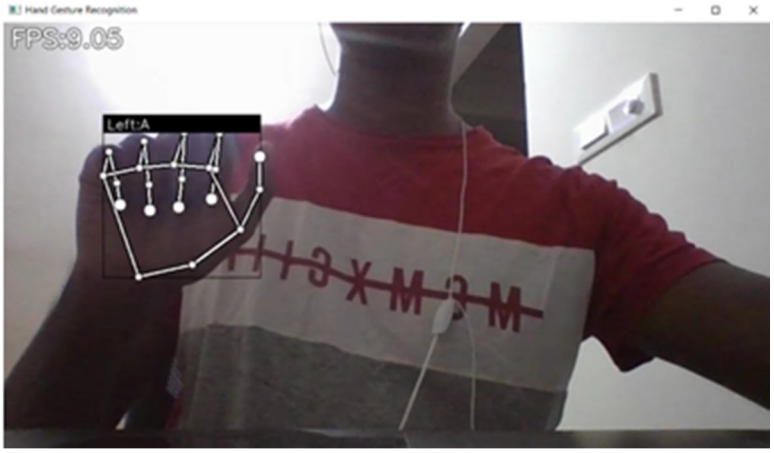



Fig 6 : ASL to English feature recognizing the alphabet 'A'

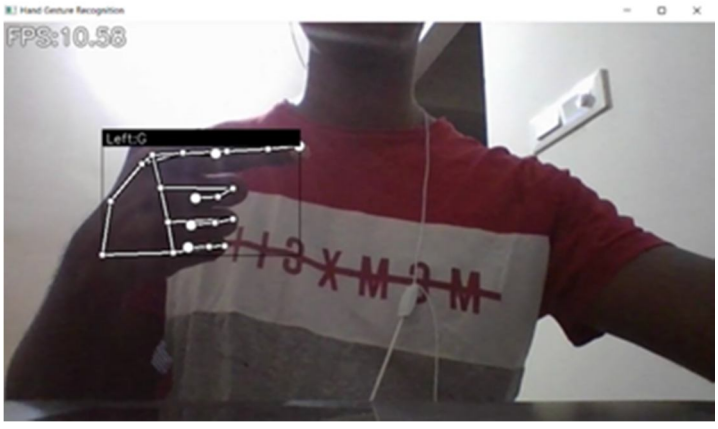



Fig 7 : ASL to English feature recognizing the alphabet 'G'

The first option provided on the home page of our application is used to convert American Sign Language gestures into the corresponding English text alphabets. This feature is targeted towards the disabled category of users. This function is facilitated by the use of Google's Machine Learning framework : MediaPipe. MediaPipe is used to build machine learning pipelines for processing time series data.

The web camera along with MediaPipe recognizes predefined index points on the palm and back of the user's hand and determines the distance between each point to learn the different hand symbols. A logging key is created and used to store the calibrated index values for each hand gesture and store it into the dataset each corresponding to the defined (trained) gesture.

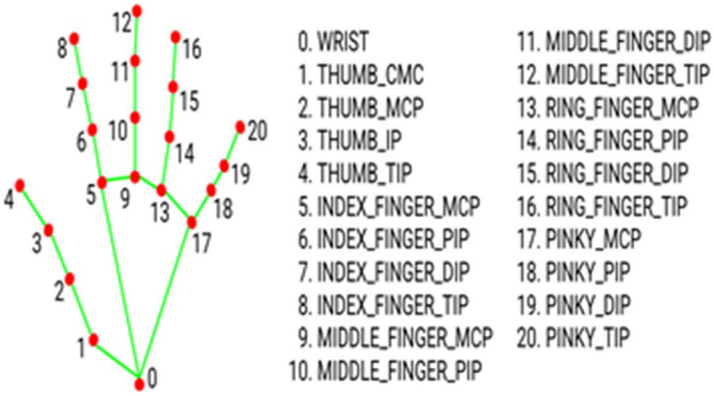



Fig 9 : MediaPipe hand landmarks

The calibrated index points that are stored into the dataset by using our defined logging key, it is stored in a tabular representation of a dataset containing only numeric values. In the first column there exist values ranging from 0 to 25 each representing an alphabet of the English script. Multiple occurrences of the same numeric constant depict different data points used to train or eventually test and validate the machine learning model. Each row there after that is appended to each constant and is used to represent values pertaining to that specific data point. The dataset that was built to implement on our machine learning model contains roughly 5900 data points of which 60 to 70% is used in the training phase while the remainder is retained for validation and testing of our model.

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0.395 | -0.09 | 0.73 | -0.34 | 0.865 | -0.665 |
| 2 | 0 | 0 | 0 | 0.47917 | -0.16146 | 0.78646 | -0.34375 | 0.92188 | -0.65625 |
| 3 | 0 | 0 | 0 | 0.38 | -0.19 | 0.655 | -0.45 | 0.78 | -0.755 |
| 4 | 0 | 0 | 0 | 0.41618 | -0.13295 | 0.72832 | -0.38728 | 0.88439 | -0.68786 |
| 5 | 0 | 0 | 0 | 0.4104 | -0.13873 | 0.75723 | -0.42197 | 0.90173 | -0.74566 |
| 6 | 0 | 0 | 0 | 0.41935 | -0.1828 | 0.7043 | -0.45699 | 0.82796 | -0.76344 |
| 7 | 0 | 0 | 0 | 0.42623 | -0.19672 | 0.69399 | -0.47541 | 0.81967 | -0.77596 |
| 8 | 0 | 0 | 0 | 0.3883 | -0.19149 | 0.66489 | -0.45745 | 0.79787 | -0.76596 |

Fig 10 : Instance of dataset depicting a few data points for the alphabet 'A'

In the above fragment instance of our dataset we can see a few of the data points and their corresponding values for the English alphabet 'A' (represented by '0').

*B. English to ASL*

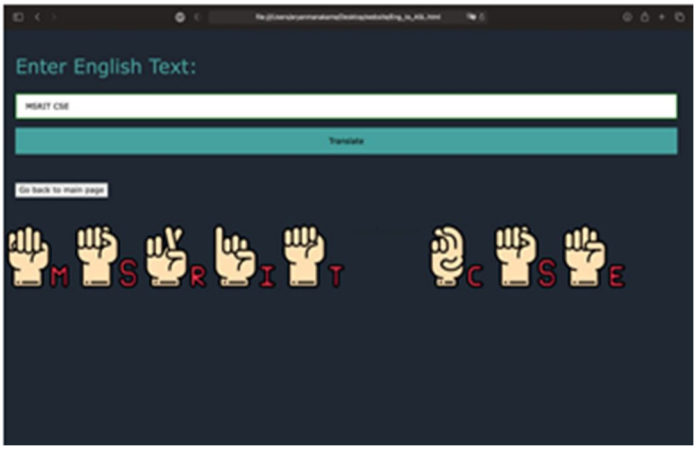



Fig 11 : 1st instance of English to ASL translation for entered English text

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

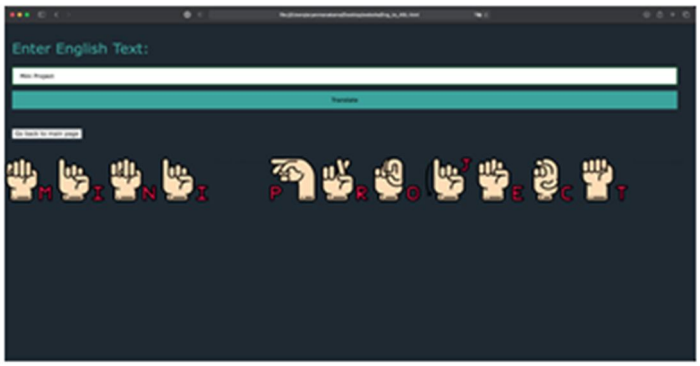



Fig 12 : 2nd instance of English to ASL translation for entered English text

The second option provided on the home page of our application is used to convert English text into the corresponding American Sign Language gestures. This feature is targeted towards the abled category of users. Through the use of HTML, CSS and Javascript I developed a web page that provides the second functionality option that is displayed to users on the home page ("English to ASL"). As and when an abled user enters a English string that is to be translated into its corresponding ASL gestures the web application performs the translation and displays the translated result. This is the second user interface page that is provided by our translation application.

## VI. CONCLUSION AND FUTURE SCOPE

### *A. Conclusion*

Our sign language translator application holds great societal value by aiming to provide better opportunities and lifestyles to those who unfortunately were not born with the same privileges as abled individuals are. The sign language translator finds its significance by identifying use cases that can be of importance to individuals who are disabled, such as job interviews between abled interviewers and disabled candidates or on focus tuition sessions, etc. The sign language translator aims to eliminate the wall that stands between abled and disabled individuals and hence is perceived as a significant and socially valuable project. Allowing and facilitating equal opportunities to all individuals of the society can in turn bring out a sense of humanity and care to enrich the social environment of that society itself.

### *B. Future Scope*

In the future as I complete developing and optimizing the sign language translation web application I aim to attempt to create a mobile application version of our translator. This mobile application will facilitate portability, easier usability and overall better reach of our translator over the internet. I also aim to facilitate the translation of standard words and phrases in order to make the application more versatile and suitable for faster communication.

## VII. ACKNOWLEDGMENT

I want to sincerely thank my mentor for helping me through the process of turning this application into a reality.

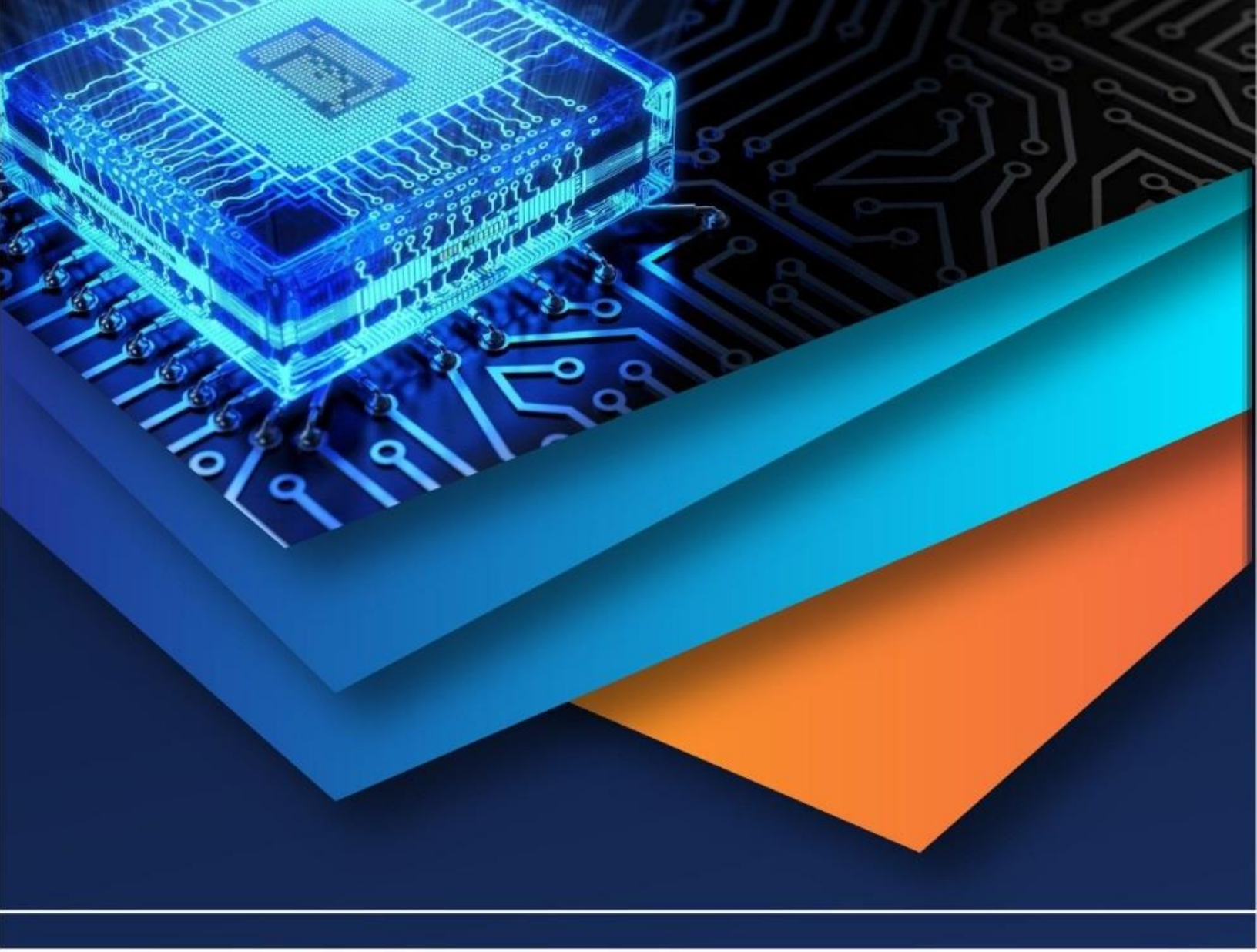
## REFERENCES

[1] Rohith Sri Sai, Mukkamala & Rella, Sindhusha & Veeravalli, Sai Nagesh. (2019). OBJECT DETECTION AND IDENTIFICATION A Project Report.
[2] Madhuri, Yellapu & Govindan, Anitha & Mariamichael, Anburajan. (2013). Vision-based sign language translation device. 2013 International Conference on Information Communication and Embedded Systems, ICICES 2013. 565-568. 10.1109/ICICES.2013.6508395.
[3] Sahoo, Ashok & Mishra, Gouri & Ravulakollu, Kiran. (2014). Sign language recognition: State of the art. ARPN Journal of Engineering and Applied Sciences. 9. 116-134.
[4] "Sign Language Recognition for Computer Vision Beginners." AnalyticsVidhya, www.analyticsvidhya.com,https://www.analyticsvidhya.com/blog/2021/06/sign-language-recognition-for-computer-vision-enthusiasts/.
[5] "Sign Language Recognition Application Systems for Deaf-Mute People: A Review Based on Input-Process-Output - ScienceDirect." Sign Language Recognition Application Systems for Deaf-Mute People: A ReviewBasedonInput-Process-OutputScienceDirect, www.sciencedirect.com, https://www.sciencedirect.com/science/article/pii/S1877050917320720
[6] Rani, Reddygari Sandhya, et al. "A Review Paper on Sign Language Recognition for The Deaf and Dumb – IJERT." A Review Paper on Sign LanguageRecognitionforTheDeaf_and_Dumb–IJERT,www.ijert.org,1Nov.2021,https://www.ijert.org/a-review-paper-on-sign-language-recognition-for-the-deaf-and-dumb.

©IJRASET: All Rights are Reserved 

[7] Tara, R.Y., P.I. Santosa, and T.B. Adji, Sign Language Recognition in Robot Teleoperation using Centroid Distance Fourier Descriptors. International Journal of Computer Applications, 2012.
[8] Oshita, M. and T. Matsunaga. Automatic learning of gesture recognition model using SOM and SVM. in International Conference on Advances in Visual Computing. 2010. Las Vegas, NV, USA: Springer-Verlag.
[9] Binh, N.D., E. Shuichi, and T. Ejima. Real-Time Hand Tracking and Gesture Recognition System. in Proceedings of International Conference on Graphics, Vision and Image 2005. Cairo - Egypt.
[10] Oka, K., Y. Sato, and H. Koike, Real-time fingertip tracking and gesture recognition. IEEE Computer Graphics and Applications, 2002. 22(6): p. 64-71.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)