International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

EXPRESSIVE HAND: Sign Language to Audio Interpreter

Sakshi Srivastava[1] , Dr. Amit kr. Tiwari[2] , Navya Srivastava[3] , Shruti Ojha[4]

1,3,4 Students, Department of Computer Science and Engineering, United Institute of Technology, Prayagraj, India

2 Head of Department, Department of Computer Science and Engineering, United Institute of Technology, Prayagraj, India

---***---

Abstract

The act of sharing or exchanging data, concepts, or emotions is known as communication. Both parties must be able to speak and comprehend the same language in order for communication to be established. It can be quite challenging for people with speech and hearing problems to communicate with the broader public. The way they communicate is through sign language. To communicate with the outside world, they are dependent on interpreters who use sign language. Being a minority, most people are unaware of the sign language they employ. Sign language communication is always necessary. Without an interpreter, communication might be challenging. One can create a machine learning-based model to get around this obstacle. It is possible to train a model to translate various sign language motions into English by recognizing them. Many individuals will find this useful while interacting and chatting with the deaf and ignorant.

INTRODUCTION

In today's world, communication is the necessary key for interaction and for explaining thoughts with others. But in the case of deaf and dumb people, the means of communication are different. Deaf is the inability to hear and dumb is the inability to speak. They communicate using sign language among themselves and with normal people but normal people are unable to understand the sign language. To bridge this gap "Sign Language to Audio Interpreter" model is build using machine learning. In order to accomplish this functionality, we developed four main modules.

- ❖ Capturing data
- ❖ Training
- ❖ Prediction
- ❖ Output Generation

The Capturing data module is used for capturing the gestures through webcam and converting them in to frames (0-27) each time. In training module machine learning model is trained with the obtained data so that it can recognize and translate sign language motions with accuracy. The Prediction module is used to predict the gestures made by the user from the gesture card. The output generation module is where the output is obtained in the form of text and audio.

LITERATURE REVIEW

We have examined various studies that have utilized different methods to translate sign languages.

- Akshatha Rani K [1] proposed a technology which is able to identify hands with varying skin tones and in various lighting scenarios, including dimly lit environments. ANN is used to classify ASL alphabet pictures.
- Sharvani Srivastava [1], Amisha Gangwar [2], proposed a system that can instantly identify letter signs in Indian Sign Language using the TensorFlow object identification API. It was an affordable solution because they used Python and OpenCV to take pictures from a webcam and used a dataset of these indicators to train the system.
- Bhavadharshini M [1] proposed system aims to create a modified ASL Translator using a CNN with YOLO (You Only Look Once) in a real-world setting. The system utilizes YOLO and CNN to detect and recognize images captured from a camera, converting them into text format for display to ordinary individuals.
- Kohsheen Tiku [1], selects the optimal approach for developing a video technology that provides deaf or mute people with text or speech translation from sign language. Their technology can recognize letters from photos with accuracy thanks to a customized SVM model.
- Dhivyasri S [1] proposed an optimum method for identifying gestures, according to the study, is to combine the SVM classifier with K-means clustering and BoV classifiers. Using this technique, they developed an intuitive app that improves the accuracy of Indian Sign Language interpretation.
- Anup Kumar [1], provides technology low-cost image-based communication assistance for those with speech impairments. More gesture recognition, improved performance in a variety of settings, and solutions for co-articulation issues are among the improvements that are required.

- Keerthi S Warrier [1], proposed a system with which smartphone cameras linked to the primary PC via 1 Wi-Fi, the setup is portable. It allows programmers to input gesture samples as specific symbols by using LabVIEW's 'Particle classification' tool to build statements and recognize movements.
- Dasari Vishal [1], there approach replaces flex sensors with IMU-equipped smart bands, reducing hardware requirements and processing overhead while improving dry electrode sign recognition.

PROBLEM STATEMENT

People with hearing impairments face significant communication barriers in their daily lives, as sign language is their primary means of expression. To bridge this gap and enhance their accessibility to the wider community, there is a pressing need for a "Sign Language to Audio Translation" System. It will help them to express their thoughts, then it will become easy for understand there feeling and thinking.

PROPOSED APPROACH

On the basis of Problem Statement, we need a platform which can overcome the gap between deaf/dumb people and normal people. There are various approaches for building this platform like application based and website based. In this paper we are going forward with the website based approach. Google TensorFlow produced the kNN Classifier that was utilized in this research. Random number calculation is necessary for the kNN classifier, but it is not easily accessible in JavaScript. In order to do this, Johannes Baagøe's "implementations of Randomness in JavaScript" work was used. Four main modules comprise our proposed approach's Sign Language to Audio Interpreter: capturing data, training, prediction, and output generation. All of the concerns and challenges listed in the Problem Statement needs to be covered in these four modules.

METHODOLOGY

1. CAPTURING DATA:

The first approach is to request for accessing the user's webcam. It creates a video element in the HTML document to initialize the webcam when access is allowed. The live video feed from the webcam is displayed using this video element as a viewport. The video element may be resized or positioned by JavaScript code to make it fit the webpage's design. The application searches especially for the initial motion that indicates the start of a sequence of gestures in the sign language. This initial gesture could be a predetermined hand gesture, posture, or pattern that the program identifies as the beginning of a sign. It initiates the proper action or event to start the capture process as soon as it detects the starting gesture. Same as the starting gesture we also have to provide the stopping gesture to stop the process. These gestures are recorded by the program in real-time, which then evaluates them for recognition and reacts appropriately. Capturing process is also involved in the training module.

2. TRAINING:

The process of teaching a machine learning more especially, the k-Nearest Neighbors (kNN) algorithm to recognize and group together various sign language motions is known as training. The kNN algorithm is loaded into the code from the 'deeplearn-knn-image-classifier' package. Next, using webcam-captured samples of sign language gestures, this algorithm is trained and initialized. A straightforward and efficient machine learning approach for classification tasks is the kNN algorithm. An input data point is compared to labeled examples in a training dataset to determine how it operates. The algorithm finds the input point's k nearest neighbors, or closest data points, and gives it the label that is most prevalent among them. To understand patterns and relationships between gestures and their meanings, the kNN algorithm is trained on samples of sign language movements. After being trained, the system can translate fresh, unknown movements into text/speech and classify them accordingly. The training dataset is stored in the "gesture card". Within the application's user interface, the gesture card functions as a visible and potentially interactive feature that offers details and feedback regarding recognized sign language gestures and their translations.

3. PREDICTION:

The kNN, machine learning technique is easy to use and efficient for categorization tasks. It operates by comparing new input (a sign language gesture in this case) with well-known examples from the training dataset. The most common class among the neighbors is chosen as the expected class for the input by the algorithm, which locates the "k" nearest examples, or nearest neighbors, to the input. It is able to identify and translate sign language movements in real-time based on their similarity to previously seen examples in the training dataset by employing the kNN algorithm for prediction.

4. OUTPUT GENERATION:

Every technology is essential to producing an output in this sign language to text/audio interpreter program. The areas are displayed below:

- Managing the application's logic and functionality, which includes:
  - Adapting the user interface dynamically in response to recognition outcomes.
  - Controlling user behaviors, like clicking buttons.
  - Working with DOM components to show gesture cards and translated text.
  - Managing the webcam to identify gestures.
- Describing the user interface's structure, comprising:

- The arrangement of components like forms, buttons, and containers.
- Including video components to enable webcam access.
- Including placeholders for gesture cards and translated text.

- Styling the HTML elements, such as:
  - Text and background colors, fonts, and sizes, to improve the application's visual presentation.
  - Arranging and placing the components.
  - Making the user experience more engaging by including animations and transitions.

Besides all this the output generated after processing is in both the for i.e. text and audio: Text Output: The application most likely produces text output on the user interface using HTML and CSS. Using the translated sign language gestures, JavaScript may also be utilized to dynamically update and change the written content. Audio Output: Its JavaScript code has the ability to cause speech synthesis events, giving users audio feedback based on motions that are translated from sign language.

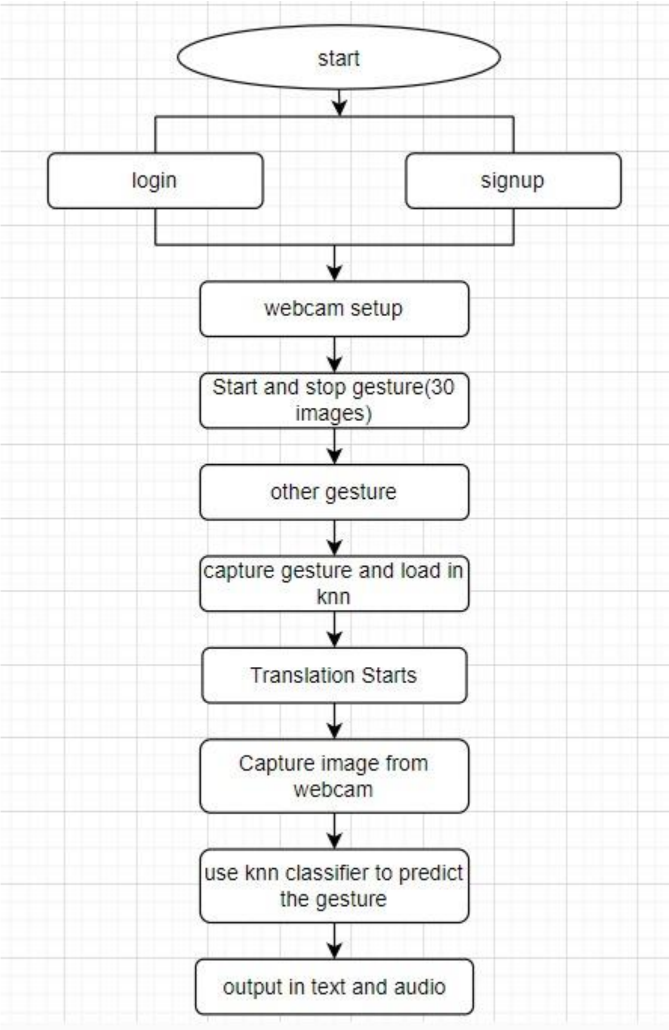

Fig. 1. Flow Chart

FUTURE SCOPE

We expect a system that can effortlessly translate any sign language into a chosen spoken language in the near future. Users won't need to supply their own training data as we improve our dataset and refine our methods for hand isolation, increasing the system's accessibility for everyone. Furthermore, it has a great deal of potential for this technology to be combined with commonly used services like video calls, or Google Duo, to improve the lives of countless people all over the world.

CONCLUSIONS

Enhancing communication between hearing and non-hearing societies is the goal of the Sign Language Translator initiative. It instantly translates sign gestures into spoken English using computer vision and tensorflow. Through data collection, model training, and the development of an understandable application, it facilitates communication.
This project can enhance education, healthcare, and social interactions by fostering inclusivity and accessibility for deaf or hard of hearing individuals. It demonstrates how innovations in computer vision and machine learning may improve people's lives by finding solutions to practical issues.

REFERENCES

1. Akshatha Rani K, Dr. N Manjanaik "Sign Language to Text-Speech Translator Using Machine Learning", international Journal of Emerging Technologies and Innovative Research, Volume 9, July 2021.

2. Sharvani Srivastava, Amisha Gangwar, Richa Mishra, Sudhakar Singh. "Sign Language Recognition System using TensorFlow Object Detection API", ANTIC, Volume 1, 5 Jan 2022.

3. Bhavadharshini M, Josephine Racheal J, Kamali M, "Sign Language Translator Using YOLO Algorithm", IOS Press in 2021.

4. Kohsheen Tiku, Jayshree Maloo, Aishwarya Ramesh, Indra R, "Realtime Conversion of Sign Language to Text and Speech", ICIRCA, 6 september 2020.IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2.

5. Dhivyasri S, Krishnaa Hari K B, Akash M, Sona M, Divyapriya S, Dr. Krishnaveni V, "An Efficient Approach for Interpretation of Indian Sign Language using Machine Learning", volume 3, ICPSC, 14 may 2021.

6. Anup Kumar, Karun Thankachan and Mevin M. Dominic, "Sign Language Recognition", rd InCI Conf. on Recent Advances in Information Technology RAIT-2016 IEEE.

7. Keerthi S Warrier, JyateenKumar Sahu, Himadri Halder, Rajkumar Koradiya, and Karthik Raj V Software Based Sign Language Converter,

International Conference on Communication and Signal Processing, April 6-8, 2016, India.

8. Dasari Vishal, H M Aishwarya, K Nishkala, B Toshitha Royan, T K Ramesh, “Sign Language to Speech Conversion Using Smart Band and Wearable Computer Technology”, 2017 IEEE International Conference on Computational Intelligence and Computing Research.