Raghwendra Kumar Pandey, 2024, 12:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology

An Open Access Journal

Hand Signs to Audio Converter

Raghwendra Kumar Pandey, Manoj Kumar Thakur, Anish Thakur, Rahul, Saurabh Agarwal, Assistant Professor Mukesh Bharadwaj

Dept. of computer science and Engineering,
Dronacharya Group of Institutions, Greater Noida

Abstract- Individuals primarily communicate with one another. Blind and deaf people use sign language to communicate with others. These individuals have difficulty communicating their message to ordinary people. Deaf and blind people believe they are unable to communicate because of a lack of communication skills, and as a result, they are unable to express their emotions. Because most individuals aren't educated in sign language, communicating in an emergency can be extremely challenging. As a consequence, the challenge may be solved by converting hand gestures into human-hearing sounds and text. Vision and non-vision approaches are two of the most commonly used methods for detecting hand movements or gestures. In a vision-based approach, a camera will be used for gesture detection, whereas sensors will be employed in a non-vision-based technique. In this study, a vision-based technique was used. This device detects and locates hand motions in order to keep a communication channel open with others. Using convolutional neural networks and artificial neural networks, this research develops a gesture recognition system. This study looks into the advantages and disadvantages of hand motion recognition.

Keywords- Speech Impaired, Communication, Speech, Hand Gesture, Image Processing

I. INTRODUCTION

In the 21st century field of science and technology, one has reached such a level that people are expecting more comfortable and useful things, which can make their lives easier. Nowadays, homes with voice recognition built in with the sense of gestures have already been conceived. There are video games in the market which can be played with real time gestures and all this has been possible with the advent of the new technology. Even our mobiles have been loaded with all similar technologies. Nevertheless, there are people who are less fortunate than us and are physically challenged, may it be deafness or being aphonic.

Figure 1: Hand Signs

And mostly from a computer scientist that computer scientists can provide some machine or a model which can help them to communicate and express their feelings with others. The deaf and the mute can only perceive visual, therefore communication is done by visual and sound. This, sign language is a medium for communication between the deaf and the mute. Sign Language. Recognition (SLR)is a tool that executes the conversion of sign language into text and further into speech. Research in SLR started two decades before, all over the world especially in American Sign Languages. Based on statistical world analysis, over 5% of the world's population – 360 million people have hearing disability. The biggest drawback for the deaf is their employment issues. Communication has always played a very vital role in getting the task solved so therefore, the issue. In order to help the deaf people communicate with

the ordinary people, we build a system to translate sign language into text and further into speech. This concept proposed a system that can automatically detect hand signs of alphabets in American Sign Language (ASL) that is all the English alphabets. To create spaces between the letters we create extra two signs, i.e., SPACE and OK.

The SPACE sign is used to create the space between different recognized words, and the OK sign is used to stop capturing and start to execute a current required function. This system is based on American Sign Language (ASL), which is considered to be a complete language. The main focus of this project is for helping the deaf and the mute by converting hand gestures to speech. Finger sign is a subset of sign language, and uses finger signs to spell words of the spoken or written language. The finger sign recognition task involves the segmentation of finger sign hand gestures from image sequences. ASL (American Sign Language)is the fourth most commonly used language in the USA and is extensively used by deaf people and this language is officially acquired by the deaf society of United States. It is a unique language that highlights signs made by moving the hands.

ASL is not defined as the world language but it has its roots in English speaking parts in Canada, few regions of Mexico, and all over United States of America. Human beings are gifted with a voice that allows them to communicate with each other. Therefore, spoken language becomes one of the main key points of humans. Unfortunately, not everybody has this capability because of one sense, i.e., hearing.

In India, there are around 5 to 15 million deaf people approx. Sign language is considered to be the basic alternative communication method between the deaf people and several dictionaries of words or single letters have been defined to make this communication strong and effective. Without an interpreter it gets difficult for such a communication to take place. Therefore, a system that converts symbols in sign languages into plain text and further into speech can help with real-time communication.

II. TECHNICAL ASPECTS

- First we create and store the hand gestures of signs with the help of image processing techniques, i.e.,converting RBG(colour) image to grayscale.
- And then again converting that grayscale image to binary image using threshold.
- Then we smooth the image using Gaussian and Median blur technique and to recognize the edges of hand gestures we use contours.
- Then we store these gestures in database and after that we use CNN algorithm on these stored images of hand gestures using Tensorflow and Keras. Tensorflow and Keras is used to test and train the system, which then recognizes the gestures which are stored in the database and gives appropriate results when the user runs the application.

III. MATERIAL AND METHOD

Machine Learning which involves the training of models and decision making based on training, nowadays can solve a wide range of real-time issues need to be addressed. The tasks that can be performed are “classification”, recognition”, “detection” followed by “predictions”. Additionally, it is very methodical to computerize the procedures that use the information in the form of data

The primitive design utilizes the available data to give as an outcome, an arbitrary model that can return various outputs. There is a chance that obtained outputs might provide an acceptable result with the novel inputs followed by offer predictions close to the existing data. Proposed model or system will facilitate the hearing and speech impaired to communicate in a better way with the entire group. As an instance, there are circumstances in which speech impaired face problems in conveying with early replier whenever there is a necessity.

Even though repliers may practice the foundations of “ASL”, it is delusional to contemplate each one to transform into a person having fluency in “ASL”. The progressions experienced in computer

recognition help the first responder in comprehension and provide support for the people who cannot communicate orally. In this work, the static ASL data has been taken . A CNN has been trained to distinguish the "signs" depicted by all the images. A Linear Model of Co-regionalization (LMC) collects the dynamic sign language dataset. However, data related to each joint of each hand are represented as (x,y,z) coordinate. Computer vision is a field of artificial intelligence that enables computers to interpret and understand visual information from the real world. It involves processing, analyzing, and extracting meaningful insights from images or videos. In hand gesture recognition, computer vision techniques are used to detect and interpret hand movements and gestures captured by cameras or other imaging devices. This involves various steps such as hand detection, tracking, feature extraction, and gesture classification. Computer vision algorithms analyze the spatial and temporal characteristics of hand movements to recognize specific gestures, which can then be translated into commands or actions for interacting with digital systems or devices.

Hand sign recognition often relies on deep learning algorithms like CNNs due to their ability to automatically learn features from images. CNNs have been successfully used in various hand sign recognition systems. Other algorithms that can be used include:

1. Convolutional Neural Networks (CNNs)
CNNs are particularly well-suited for image recognition tasks like hand sign recognition due to their ability to learn hierarchical features directly from pixel values.

2. Networks (RNNs)
RNNs can be used for sequential hand sign recognition tasks, where the temporal information of hand movements is important.

3. Deep Belief Networks (DBNs)
DBNs are generative models that can be used for feature learning and classification tasks. They have been applied to hand sign recognition with promising results.

4. Support Vector Machines (SVMs)
SVMs can be used for hand sign recognition by transforming the image data into a higher-dimensional space where a hyper plane can separate different classes of hand signs.

5. K-Nearest Neighbors (KNN)
KNN is a simple and effective algorithm for hand sign recognition, especially in situations where the dataset is not very large.

6. Decision Trees and Random Forests
These algorithms can be used for hand sign recognition by recursively partitioning the feature space based on hand sign attributes.

"Each of these algorithms has its strengths and weaknesses, and the choice depends on factors such as the complexity of the task, size of the dataset, and computational resources available. For hand sign recognition, CNNs are often preferred due to their high performance and ability to automatically learn relevant features from raw image data."

IV. SIGN LANGUAGE REPRESENTATION CAN BE APPROACHED IN VARIOUS WAYS

1. Hand Posture Classification
This approach involves recognizing hand postures or configurations to interpret sign language. It typically relies on computer vision techniques to extract features from hand images and then classify them into predefined hand postures corresponding to specific signs.

2. Skeletal Hand Tracking
Skeletal hand tracking involves capturing the movements of the hand joints in real-time using depth sensors or cameras. This approach provides dynamic information about hand movements, which is essential for interpreting sign language gestures that involve motion.

3. Gesture Recognition
Gesture recognition involves recognizing and interpreting sequences of hand movements or

gestures to understand sign language. It can combine techniques from both computer vision and machine learning to analyze temporal patterns in hand movements.

4. Deep Learning Models

Deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), can be applied to sign language representation tasks. CNNs are well-suited for static hand posture classification, while RNNs are effective for capturing temporal dependencies in gesture sequences.

5. 3D Hand Pose Estimation

This approach involves estimating the 3D positions of hand joints from 2D images or depth data. It provides more detailed spatial information about hand movements compared to 2D methods and can be useful for accurate sign language representation.

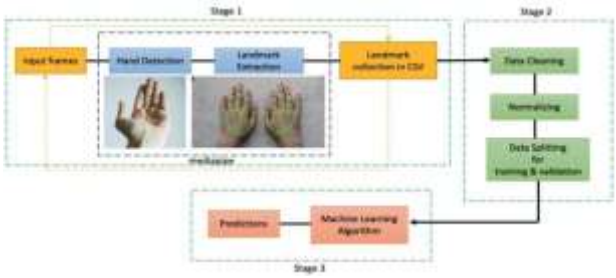

Figure 2: 3D Hand Pose

6. Data Gloves and Wearable Sensors

Wearable sensors, such as data gloves equipped with inertial sensors or flex sensors, can capture hand movements directly. These sensors provide precise information about hand gestures and can be used for real-time sign language interpretation.

7. Combined Approaches

Some systems combine multiple approaches, such as using hand posture classification for static signs and skeletal hand tracking for dynamic signs. This hybrid approach leverages the strengths of each method to improve overall accuracy and robustness.

"Each approach has its own set of advantages and limitations, and the choice depends on factors such as the application requirements, available technology, and the specific characteristics of the sign language being represented."

V. CONCLUSION

Sign language may be a helpful gizmo to ease the communication between the speech impaired community and additionally the normal people. As speech impaired use their standard sign language, it cannot be easily understandable by common people, and the visually impaired cannot see their gestures. This project aims to lower the communication gap between this community and additionally the standard world. The projected methodology interprets language into speech. The system overcomes the necessary time difficulties of speech-impaired people and improves their manner. This method transforms the language in linking flow speech that is very much comprehensible by visually impaired and normal people. In world applications, this system is helpful for voice impaired of us who cannot communicate with normal persons. This system provides us with a high gesture recognition rate with an accuracy of 95% within minimum time and is getting converted into speech with an accuracy of 96%. It is also functional for speech impaired, immobilized inhabitants, and for Smart Home and commercial applications.

REFERENCES

1. Eshetu, Yigremachew, and Endashaw Wolde. "A real-time Ethiopian sign language to audio converter." International Journal of Engineering Research and technology (IJERT) 8
2. Eshetu, Y., & Wolde, E. (2019). A real-time Ethiopian sign language to audio converter. International Journal of Engineering Research and technology (IJERT), 8, 661-66.
3. Eshetu Y, Wolde E. A real-time Ethiopian sign language to audio converter. International Journal of Engineering Research and technology (IJERT). 2019;8:661-.

4. Nagpal, Aastha, et al. "Hand sign translation to audio message and text message: A device." 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN). IEEE, 2020.
5. Nagpal, Aastha, Ketaki Singha, Rakshita Gouri, Aqusa Noor, and Ashish Bagwari. "Hand sign translation to audio message and text message: A device." In 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), pp. 243-245. IEEE, 2020.
6. Nagpal A, Singha K, Gouri R, Noor A, Bagwari A. Hand sign translation to audio message and text message: A device. In2020 12th International Conference on Computational Intelligence and Communication Networks (CICN) 2020 Sep 25 (pp. 243-245). IEEE.
7. Gayitri, H. M. (2022). Hand Gesture Recognition and Voice Conversion for Speech Impaired. Mathematical Statistician and Engineering Applications, 71(4), 2739-2746.
8. Gayitri , H. M. "Hand Gesture Recognition and Voice Conversion for Speech Impaired." Mathematical Statistician and Engineering Applications 71.4 (2022): 2739-2746.