DNN BASED GESTURE LANGUAGE TRANSLATOR

E.RAMASSAMY[1], S.THAMIZHARUVI[2], V.KARTHIKAVENGAT[3]

[1]Assistant Professor, Dept of ECE, Manakula Vinayagar Institute of Technology, Puducherry.

[2,3]Final Year Students, Dept of ECE, Manakula Vinayagar Institute of Technology, Puducherry.

ABSTRACT

Gesture recognition is a powerful tool for creating interfaces that are adaptable to the needs of users. However, recognizing gestures can be challenging due to the inherent variability of human gestures. One application of gesture recognition is in sign language, which is used for communication and interaction between people. There are many different systems and methods for recognizing sign language, and our approach focuses on recognizing 26 symbols of hand motions used in the American Sign Language (ASL) for gesture recognition. The system is both sturdy and effective, allowing for real-time identification of static images classification by MK-ROD algorithm is proposed. The core idea behind a DNN-based gesture language translator is to use a sequence of input frames of a person making hand gestures and translate them into spoken language or written text. The procedure entails the utilization of a vast collection of annotated pairs of gestures and language to train a DNN model, which allows the model to acquire the ability to associate the two modalities. A success of DNN-based gesture language translators heavily relies on the quality, diversity of the training data and providing accuracy.

KEYWORDS*: Raspberry-pi, Camera, Speaker*

INTRODUCTION

A gesture language translator refers to a technological solution that employs computer vision and artificial intelligence to detect and comprehend human gestures and render them into spoken or written language. It develops a communication bridge- between differently abled people and who do not know the sign language. Gesture language translators usually utilize sensors or cameras to capture the gestures of the user, which are then processed through machine learning algorithms to interpret and translate them into speech or text.

These translators have broad applications in various settings, including hospitals, workplaces, public spaces, and educational institutions, and are designed to promote inclusion and communication for individuals with hearing impairments or deafness. The potential of gesture language translators to transform the way we communicate and eliminate obstacles for sign language users is enormous. They have the ability to enhance communication and promote inclusivity for individuals who depend on the sign language as their basic mode to communicate. They can provide more efficient and accurate communication, improve accessibility and inclusion, and promote greater understanding and respect for diversity.

LITERATURE SURVEY

The present hand gesture recognition system is affected by its surrounding environments. This led to misclassification of hand gestures for automotive functioning system. The hand gesture can be detected and recognition under any background environment. Hence, there is a requirement for developing a hand gesture recognition system under different environment. Gesture is one of the sign language representation methods which find many applications in real time world. HCI (Human Computer Interaction) is a method generated for interacting the human attitude with computer system to produce a user interface for easy interfacing between human and computer [1].

Various gestures of human were incorporated with computer system and made to recognize by the system, a few gestures like hand gestures, head movements, facial expressions, lip reading, etc. were considered. A gesture may be accurately defined as an expression using a body movement with vocal communication or without vocal communication. The gesture may be done through any part of the body as with hands, face, palm, eye, etc [2].

Gestures involve moving parts of the human body to convey information or commands. Gestures are physical movements of the hands, arms, or face that are intended to convey the information by the differently abled person. Identifying gestures entails tracking human body movements and comprehending their significance as specific commands or information [3].

The normal gesture research network has recognized the basic kinds of regular Hand Gestures (HG) acknowledgment methods. Common HG correspondence is an exceptionally dynamic exploration territory in the field of Computer Vision. Such a correspondence strategy furnishes the straightforwardness to connect with machines without the necessity of any additional gadget. Also, if the clients don't have adequate specialized information about the framework, they are as yet ready to utilize the framework with their uncovered hands. When conveying, people frequently express their announcements through gestures [4].

The gestures are delineated normally with the words and the gestures help upgrade the comprehension (of the correspondence) of the audience. Gestures permit people to impart sentiments and contemplations with various feelings with words or without words (Shin & Lee 2006). A human gesture can be by any means; however, few have an extraordinary importance.In modern robotic world, the interaction between human and computer is important to operate the robots in an automatic manner. In past decades, the hand gesture (a sign language) is used to interact with machine. Hence, the recognition of hand gesture by a machine is essential for performing a particular task. The hand gestures can be detected by gloves, optical markers and image processing techniques in conventional methods [5].

These methods are having some limitations and hence they are not providing optimum hand gesture recognition rate. In case of glove technique, many sensors are placed within the glove region and they detect the position and orientation of hand in an effective manner.

Though this method is more effective, it is highly expensive and it requires wired communication between gloves and system or machine. In case of optical marker method, InfraRed (IR) is used for detecting the location and orientation of moving hand. This method is also highly expensive and the design process is more complex than the glove-based hand gesture recognition system [6].

The suggested system for gesture recognition employs solely the deep information to detect unrestricted symbols. This process consists of three process the first stage is pre-processing, static hand posture recognition, segmentation and feature extraction. These static hand postures, in conjunction with the hand's trajectory, are then prepared to identify the moving actions and symbols by the hand [7].

This process not only recognize the symbols also the back-ground. A system assumes that the primary user is in close proximity to the camera and occupies a significant portion of its field of view, enabling the extraction of the user's body as the most extensive area within 1000 to 1850 milli meters. Background subtraction can be used to eliminate invalid gestures that are not produced by the user. According to the datasheet, the depth camera is capable of providing the depth of objects within a range of 810 to 4000 millimeters. To detect hand gestures, the proposed system requires a hand region exceeding 1600 pixels; hence the user must remain within 1000 to 1850 millimeters. Both hand detection and tracking are implemented in the prototype system [8].

To describe the dynamic hand gestures accurately, the system utilizes the hand positions and degrees that signify the transition. The proposed approach initially captures the feature descriptors of the hand and then categorizes the hand into possible postures, as depicted in Figure 1. The system designates the contour of the hand, including the first dorsal interosseous, between the fingertips of the thumb and the index finger, as the region of interest for feature extraction (RoF).The segmentation process divides an image into two parts: the background and the foreground, which includes the hand region, i.e., the region of interest. The resulting segmented image assigns the pixel value '1' to the hand region and '0' to the background. This image acts as a mask and is used to extract the hand region from the RGB image by performing an element-wise multiplication of the binary image with the original RGB image. The image is then resized to reduce the matrix size used for recognition. The images are transformed into column matrices to facilitate feature extraction [9].

Gesture recognition is a powerful tool for creating interfaces that are adaptable to the needs of users. However, recognizing gestures can be challenging due to the inherent variability of human gestures. One application of gesture recognition is in sign language, which is used for communication and

interaction between people. There are two main approaches to hand locating and sign language recognition: traditional methods and deep learning methods. In recent years, deep learning has shown great promise in computer vision, offering advantages such as rich feature extraction, strong modeling ability, and intuitive training. This paper proposes a neural network-based approach to hand locating and sign language recognition for common sign language in India, with a focus on recognizing Indian Sign Language and converting it into speech and text in two languages: English and Malayalam. The proposed system aims to provide a more convenient and accessible means of communication for differently abled person [10].

EXISTING TECHNOLOGY

The process divides an image into two parts: the background and the foreground, which includes the hand region, i.e., the region of interest. The resulting segmented image assigns the pixel value '1' to the hand region and '0' to the background. This image acts as a mask and is used to extract the hand region from the RGB image by performing an element-wise multiplication of the binary image with the original RGB image.

The image is then resized to reduce the matrix size used for recognition. The images are transformed into column matrices to facilitate feature extraction. Feature extraction is a critical step in the recognition process since it reduces the data dimensionality. In this stage, we extract features using Principal Component Analysis (PCA), which provides us with Eigen Values and Eigen Vectors. To calculate these Eigen Values and Eigen Vectors, we first create a column matrix of all the images and then concatenate them to form a single matrix. This process facilitates dimensionality reduction and improves the efficiency of the system.

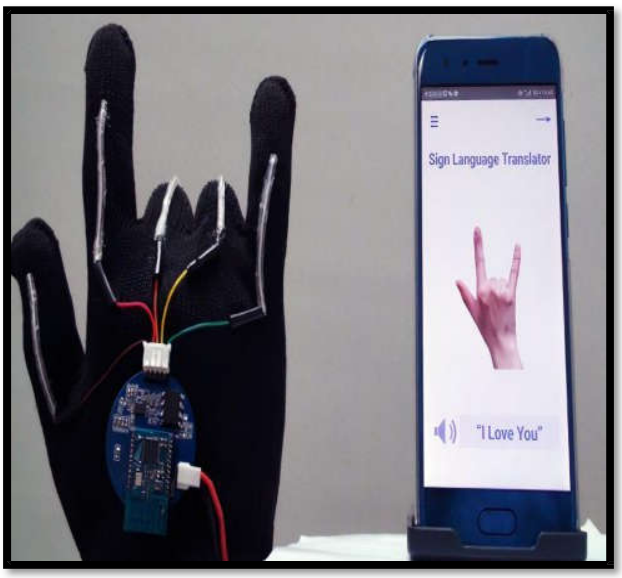

Figure.1: Existing Technology of Gesture Language Translator

PROPOSED SYSTEM

1. PRELUDE

The proposed system presented in this paper involves two stages, object detection and classification. Hand detection is done using the Single Shot Multi Box Detection (SSD) architecture, while a deep learning structure based on the Inception v3 plus Support Vector Machine (SVM) is proposed for feature extraction and classification of hand gestures. The system is designed using a sign language fingerspelling dataset, and the results show the effectiveness of the proposed hybrid structure in sign language translation.This process converts the video process into natural process in the existing models typically represent sign visual features in a frame-wise manner, which neglects the temporal information of signs and leads to translation ambiguity.

This paper proposes a method that utilizes the temporal semantic structures of sign videos to learn more discriminative features. The paper mainly discusses image or vision-based SLR systems that involve feature extraction and classification and briefly describes the translation of SL to speech. Overall, this paper aims to provide acomprehensive introduction to automatic hand gesture recognition and sign language interpretation

The main purpose of the project is to facilitate communication between individuals who use different forms of sign language or who use sign language and spoken language A gesture language translator typically works by using cameras to capture the movements of the signer's hands and body, and then processing this information through machine learning algorithms to produce a spoken or written translation.

2. FLOW CHART

A data flow diagram (DFD) is a visual representation that illustrates how information moves through a system, including input data, processing activities, and output data. It provides a graphical depiction of the flow of data and the transformations applied as it moves through the system.The bubble chart is one of the most precious tool for modelling in data flow diagram is used for making components in the system and the data process.In this process input is given by the user and the system gives the output. This process generalizes the technique and gives the detailed information.

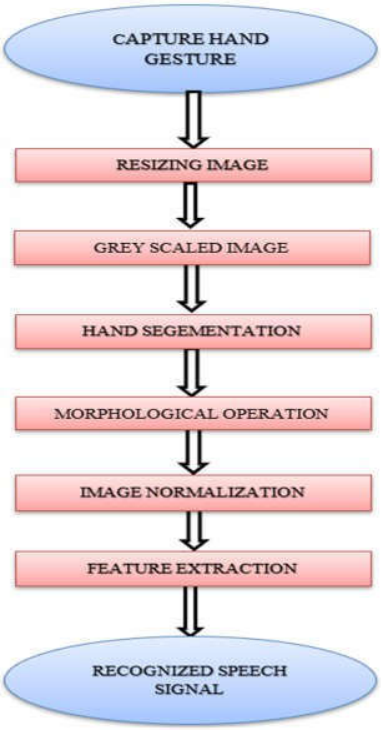

Figure.2: Flow Chart of DNN based Gesture Language Translator

3. BLOCK DIAGRAM

The step of Gesture Recognition typically involves several stages, starting with Data Acquisition, which involves capturing data in the form of hand gestures using a camera or motion sensor.

The next stage is Pre-processing, where acquired data is prepared by removing any noise or unwanted information from the input signal. This may include steps such as filtering, smoothing, and normalization .The following stage is Feature Extraction, where relevant features such as the shape, size, and movement of the hand are extracted from the pre-processed data. The process of recognizing hand gestures involves using the extracted features to identify the specific gesture.

The output is the final stage being performed involves outputting the translated speech to the user, typically through a speaker or headphones. This task is performed by Deep Neural Network algorithm and machine learning process. System or a pre-recorded database of spoken words or phrases. The hand gesture has been recognized, the system can translate it into spoken language using a text-to-speech (TTS). This task is performed by Deep Neural Network algorithm and machine learning process. system or a pre-recorded database of spoken words or phrases. The hand gesture has been recognized, the system can translate it into spoken language using a text-to-speech (TTS)

Gesture Language Translator is a complex system that involves multiple stages of data processing and analysis. By leveraging the power of machine learning techniques, these systems are able to identify the gestures and convert them into language spoken by others with great number of accuracy. It is also known as circle chart.

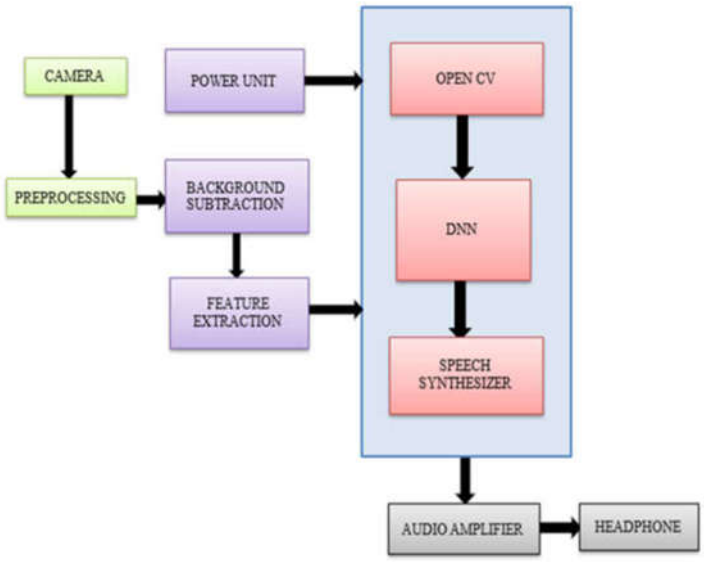

Figure.3: Block diagram of DNN based Gesture Language Translator

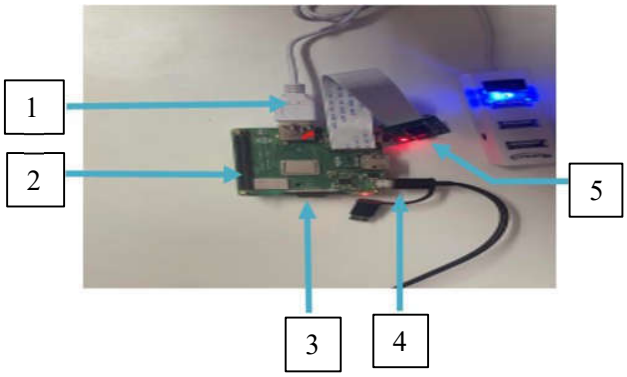

Figure.4: Hardware of DNN based Gesture Language Translator

4. HARDWARE REQUIREMENT

- Power supply
- Raspberry Pi
- Memory card
- Headphones
- Raspberry Pi camera

5. MERITS OF PROPOSED WORK

- Improved communication
- Accessibility
- Convenience
- Cost-effective
- Customizable
- Educational tool

6. WEB APPLICATIONS

The web application is the software program is accessed and used over the internet through a web browser. Unlike traditional desktop applications, web apps are not installed on a local computer, but instead are hosted on a server and accessed remotely. Web applications are designed to be accessed by users through a web browser, making them easily accessible from any device with an internet connection. They can be used for a variety of purposes, such as online shopping, social networking, online banking, emails.

a. Mobile showing a gesture
b. Raspberry pi-3 camera.

7. OUTLINE AND SUMMARY

In conclusion, gesture language translator technology has the potential to be a game-changer for individuals who cannot speak, hear as well as pupils who are not good in sign language. With advances in machine learning and computer vision, gesture recognition systems can accurately translate hand gestures into spoken language, text even sign language .This technology has the potential to bridge communication gaps and promote inclusivity in a variety of settings, including schools, workplaces, and public spaces.

However, it is important to note that gesture language translator technology is the starting stages to develop and there will be accuracy and effectiveness. A gesture language translator is a portable device that uses a camera and a Raspberry Pi 3 computer to interpret and translate sign language gestures into spoken language.

The device can be customized to recognize specific gestures or a sign used in a particular region or language, and it offers improved communication and accessibility for individuals with hearing impairments. VNC Viewer software can also be used to remotely control and access the gesture language translator's graphical user interface making it convenient and easy to use in a variety of settings. Overall, the gesture language translator offers a cost-effective, customizable, and educational solution for individuals with hearing impairments, enhancing their communication and accessibility in various environments.

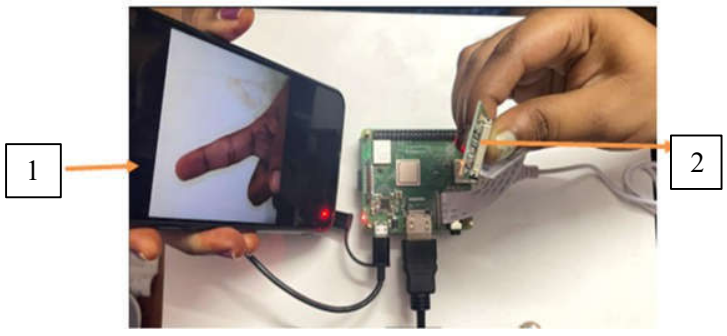

Figure.5: Working of DNN based Gesture Language Translator

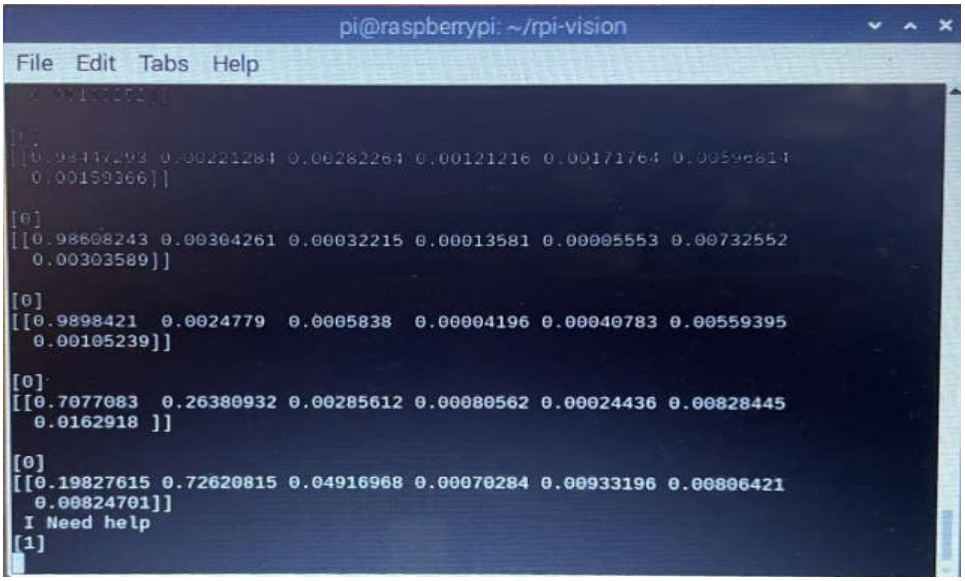

Figure.6: Output of DNN based Gesture Language Translator

8. FUTURE SCOPE

In simple words, the future scope of gesture language translator technology is very promising. This technology has the potential to improve communication and accessibility who cannot speak , hear as well as pupils who not fluent in sign language.Some of the potential future developments in this field could include improved accuracy, integration with other technologies, miniaturization, real-time translation, and multi-language support. These developments could make gesture language translator technology more accurate, portable, and useful in a variety of settings. Ultimately, this technology could help to create a more inclusive and accessible world for individuals with hearing or speech impairments.

REFERENCES

[1] He, Siming. "Research of a sign language translation system based on deep learning." In 2019 International conference on artificial intelligence and advanced manufacturing (AIAM), pp. 392-396. IEEE, 2019.

[2] Ginosar, Shiry, Amir Bar, Gefen Kohavi, Caroline Chan, Andrew Owens, and Jitendra Malik. "Learning individual styles of conversational gesture." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3497-3506. 2019.

[3] Bragg, Danielle, Oscar Koller, Mary Bellard, LarwanBerke, Patrick Boudreault, Annelies Braffort, Naomi Caselli et al. "Sign language recognition, generation, and translation: An interdisciplinary perspective." In Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility, pp. 16-31. 2019.

[4] Artemov, M. D., Lilia I. Voronova, and A. G. Vovik. "Designing Soft-Hardware Complex for Gesture Language Recognition using Neural Network Methods." In *2020* International Conference on Engineering Management of Communication and Technology (EMCTECH), pp. 1-6. IEEE, 2020

[5] Farooq, Uzma, Mohd Shafry Mohd Rahim, Nabeel Sabir, Amir Hussain, and Adnan Abid. "Advances in machine translation for sign language: approaches, limitations, and challenges." Neural Computing and Applications 33, no. 21 (2021): 14357-14399.

[6] Ardiansyah, Andra, Brandon Hitoyoshi, Mario Halim, Novita Hanafiah, and Aswin Wibisurya. "Systematic literature review: American sign language translator." Procedia Computer Science 179 (2021): 541-549.

[7] Zhang, Qian, JiaZhen Jing, Dong Wang, and Run Zhao. "Wearsign: Pushing the limit of sign language translation using inertial and EMG wearables." Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 6, no. 1 (2022): 1-27.

[8] Shah, Kashish, Sanket Rathi, Rishabh Shetty, and Kamal Mistry. "A comprehensive review on text to Indian Sign Language Translation systems." Smart Trends in Computing and Communications: Proceedings of SmartCom 2021 (2022): 505-513.

[9] Moore, Kayleigh, Stefano Pecile, and Mahdi Yazdanpour. "Gesture-based American Sign Language (ASL) Translation System." (2023).

Pandey, Surya. "Automated Gesture Recognition and Speech Conversion Tool for Speech Impaired." In Proceedings of Third International Conference on Advances in Computer Engineering and Communication Systems: ICACECS 2022, pp. 467-476. Singapore: Springer Nature Singapore, 2023.