Research and Reviews: Advancement in Robotics
Volume 5 Issue 3

MAKATON- Sign Language Recognition

Divya K K , Mohammed Bilal Afrad, Hassan Abid, Ismail Faizan, Ahammed Mehfooz Farooq
Department of Computer Science and Engineering
P A College of Engineering, Mangalore, India.

***Corresponding Author**
E-mail Id:-*divujithesh@gmail.com*

ABSTRACT
One of the standard sign language techniques is the use of hand gestures. The world is exceedingly tough to converse with for those who are hard of hearing. This project offers a solution that will allow disabled individuals to effortlessly communicate with sighted people by not only automatically recognising hand gestures but also converting them into text and vocal output. A computer-attached camera will take pictures of hands, and contour feature extraction will be utilised to identify the person's hand gestures. based on the acknowledged

Keywords: *Raspberry Pi, convolutional neural networks (CNN), OpenCV*

INTRODUCTION

A true disability is the inability to communicate. This disorder affects how people interact with one another. They have access to a wide range of communication channels, sign language being one of the more common ones. Human body language and sign language are two ways that people can communicate. Each word is made up of a variety of human gestures that signify a certain expression. Using an understanding of human motions, this approach aims to translate human sign language into voice.

To do this, we employ the web camera and speaker of the Raspberry Pi. A few systems exist that translate sign language to voice, but none of them offer a portable user interface. For instance, if a person with a speech handicap can perform in front of the system, and if the system converts the human motions into speech and amplifies them, the person may really be able to speak to a big audience.

Communication between visually and verbally handicapped people is also made possible by the device.

Human-machine contact is becoming more and more necessary as computer technology develops. Although the use of touch-screen technology is influenced by our cell phones, it is still too expensive to be adopted in professional settings. Although the mouse is very helpful for controlling devices, those who are physically challenged or unaccustomed with using the mouse for connection may find it challenging to utilise.

As computer technology advances, human-machine interaction is becoming more and more essential. The use of touch-screen technology is influenced by our cell phones, yet it is still too pricey to be used in professional contexts. The mouse is an extremely useful tool for controlling devices, but it might be difficult to use for people who are physically weak or are not accustomed to using the mouse for connections.

The research study investigates letter recognition using a hand language derived from American Sign Language. The detecting procedure makes use of contour analysis and feature extraction. The use of numerous computer vision methods and algorithms to identify hand gestures is discussed in the study.

OBJECTIVE

The goal of our research is to reduce the disparity between the general population and individuals with speech and hearing problems. The major objective of this research is to develop a system that enables deaf people to effectively communicate with others by using their natural gestures.

PROPOSED SYSTEM

One of the main issues with the present system is its lack of mobility. People used to create sign languages while seated in front of the machine. With the use of our gear, people may now utilise sign language to converse with their opponents whenever they want and from any location. We are, in other words, modifying technology to be portable. Accuracy is crucial in all systems, as was previously stated for the current system. We employ a programme called Yolo on our devices to increase the accuracy of the video that is recorded. In order to record people's hand actions, we're also using a Raspberry Pi camera. This study makes use of convolutional neural networks (CNN).

Convolutional neural networks have made one of the biggest contributions to the field of computer vision. CNNs are deep learning algorithms that make use of a feed-forward ANN. The voice is produced using "pyttsx," a piece of software. The employed programming language is Python.

LITERATURE SURVEY

Contour detection and the fuzzy c-means algorithm were used to construct a portable real-time sign language recognition system by Dr. Gomathi V. and Muthu Mariappan H. Utilizing the contours of the face, left, and right hands, they are found.

A method for deciphering Indian sign language from live video was proposed by Joyeeta Singha and Karen Das. The dataset had 480 images of 24 ISL signs with the signatures of 20 individuals. The system tested on 20 videos and achieved an accuracy rating of 96.25 percent.

An image processing, deep learning, and computer vision-based real-time two-way sign language communication system was proposed by Tanuj Bohra et al. The CNN model was able to correctly identify 17600 test photographs in 14 seconds after being trained with a sizable dataset spanning 40 classes.

Mahesh Kumar created a technique based on linear discriminant analysis that can recognise 26 hand gestures used in Indian sign language (LDA). Pre-processing methods for the dataset include morphological procedures and skin segmentation. The skin was divided up using the Otsu algorithm.

FEASIBILITY STUDY

Researchers will examine a number of hand gesture and sign language recognition techniques that have been put forth in the past as part of this project. For the dumb and the deaf, sign language is their sole form of communication. The sign language used by these people with physical disabilities helps them express their feelings and ideas to others. This paper's discussion of the numerous deaf-mute communication translator systems in use today serves as its main goal.

Deaf-mute people use two major categories of communication methodologies: Wearable Communication Devices and Deaf-Mute Communication Techniques. Examples of wearable communication devices include keypad systems, glove-based systems, and the Handicom.

In all three of the above-mentioned sub-divided approaches, various sensors, an accelerometer, a suitable microcontroller, a text- to-speech conversion module, a keypad, and a touch-screen are utilised. It is feasible to do away with the need for an external device to translate communications between deaf-mute and hearing persons. A feasibility study evaluates the likelihood of a project's success. Therefore, the study's reputation among potential investors and lenders depends heavily on how objective it is regarded to be. As a result, in order to provide data from which judgements can be made, it must be carried out impartially and objectively. Three crucial feasibility studies that are necessary for our proposal will be covered.

Operational Feasibility is a statistic that assesses how well a proposed system addresses users' issues. The project's human resources decide the operational feasibility, which comprises determining whether the system will be used once it has been installed and created. Users will find the project useful because practically all teachers and staff are already familiar with digital technologies.

Economic feasibility: This concept pertains to the question of whether the anticipated profit will be larger than or equal to the anticipated cost. The term "cost-benefit analysis" or "cost-benefit analysis" may also be used. The method involves assessing the system's anticipated benefits and savings and contrasting them with the costs. The suggested system is anticipated to provide advantages greater than its costs. With no development expenditures, this is a low-budget project. Utilizing and understanding the system is straightforward. Because of this, there is no requirement that users spend money on training. By incorporating features for both teachers and students, this system has the potential to grow. The venture might therefore later have financial success.

Technical Feasibility: A project's technical viability in terms of resources (such as software, hardware, labour, and knowledge) to accomplish it is assessed. It investigates how the resources of the proposed system will be chosen. Because it was created in Python, the system is cross-platform. Users of the system can therefore operate on any platform and have average processing capabilities.

METHODOLOGY

The human hand is a flexible entity that can deform to many different shapes. The input image must go through numerous steps in order to identify sign language, through which the necessary data will be gathered. Each procedure involves a number of different ways. An image can be captured with a fixed camera or by employing many cameras, which is the common method. The image must first be obtained and then segmented so that only the necessary data is gathered. After segmenting the image, the segmented region will be used to extract the elements needed to identify the gesture. The learning algorithms were then used to study the retrieved image. Data science approaches are used in gesture recognition, including model building, model training and testing, and data classification. There were a number of challenges encountered while putting the algorithms and approaches into practise, however these issues can be successfully fixed.

SYSTEM REQUIREMENT

Software Requirement, also known as Product Requirements Specification, is a document that describes the characteristics of a project, piece of software, or software application. A project handbook that is written before the commencement of a project or application is what a software requirement document is, to put it simply. Additional titles for this document are software document and SRS report. A software document is produced for a project, programme, or other kind of application. There are certain rules to follow when creating the software requirement specification document. The project's objective, scope, functional and non-functional requirements, and hardware and software requirements are all covered in this section. It also offers details on, among other things, the project's environmental needs, safety and security specifications, and software quality issues.

SUPPORTABILITY

Use is made of the Logi-Tech web camera, which is attached to the Raspberry PI system and installed on top of the portable device. The suggested system makes use of a camera with a resolution of 16 mega pixels, USB, and night vision. At any time of day or night, the camera can scan for motions, and the expense is negligible. Sending the scanned images to the Raspberry PI allows it to play back audio through speakers. The raspberry PI and speaker are connected, and the speaker will play the output as speech The method that is being proposed uses a straightforward speaker that is only used for audio. A gadget that is connected to a mobile device has the suggested system installed on it. The device's top-mounted camera doesn't rely ontaking pictures at a precise angle because it is located there. For the image to be captured by this system, the user must place their hands in front of the camera. The libraries and modules of OpenCV were used to construct the proposed system. They produce results immediately and are incredibly accurate and efficient.

DESIGN CONSTRAINTS

Before making any hand motions in front of the camera, thehand is first brought in front of the device. The voice is then broadcast through the speakers when the camera recognises the movements. The study's findings show that the proposed system has a quick processing time and a high accuracy rate. A gadgetthat is connected to a mobile device has the suggested system installed on it. The device's top-mounted camera doesn't rely on taking pictures at a precise angle because it is located there. For the image to be captured by this system, the user must place their hands in front of the camera. The suggested system, called MAKATON, is composed of OpenCV libraries and modules. They produce results immediately and are incredibly accurate and efficient. They contain the critical areas with distinguishing characteristics needed to learn and anticipate gestures, as shown. Furthermore, we only keep the crucial portions of the gestures in the dataset rather than the entire gesture because storing the entire gesture may reduce forecasting speed, accuracy, and efficiency.

INTERFACES

Image to text interface translation in user interfaces Everyone was taken into consideration when designing the user interface. The user interface should be straightforward and easy to use as a result. In this interface, the user's movements are recognised and shown by a panel at the bottom of the screen, and a text box converts those movements intotyped text.

Translation from text to image interface Everyone was taken into consideration when designing the user interface. You can type text in the text field of this user interface. The text in this text section is translated into sign language on a screen at the top.

Equipment Interfaces: The machine will have 64-bit architecture on theserver side.
Any personal computer that can run Windows with a mouse or that can interact with Android smartphones is allowed on

the client-side.

Software interfaces: Software is a group of programmes that instruct a computer to carry out a certain task when instructed to do so. These apps are made by programmers in order to communicate with the hardware of the system. Python 3.6 is the scripting language, while Raspbian is the operating system. The suggested system requires two pieces of software:

Server-side: The generated application will be accessed using a web browser. There will be constructed a database with about 20 sign language entries.

Client-Side: A web browser with at least HTML 3.2 support, java support, a camera, and the ability to run on any operating system that supports browsers is necessary.

Transmission Control Protocol/Internet Protocol, or TCP/IP, is the industry-standard protocol for transmitting data between a server and a client. At the highest level of communication between the web server and the client, the Hyper Text Transfer Protocol (HTTP, default port=80), will be used.

DESIGN PART

A solution to an issue described in the requirements document is what the design seeks to do. The process of moving from the problem domain to the solution domain begins with this. To put it another way, designing from a place of need enables us to choose how to satisfy those requirements. Software quality is most likely determined by the system's architecture, which also has a big impact on subsequent phases, especially testing and maintenance.

The objective of sign language recognition (SLR) systems is to give hearing aid users a quick and precise way to convert sign language into text or voice.

Determining the modules that should be included in the system, their specifications, and how they interact to produce the desired outputs are the main objectives of system design. At the conclusion of the system design, all significant data structures, file formats, output formats, and major modules are chosen, along with their requirements.

System Structure is concerned with the structure and behaviour of the system as established by the conceptual design. A comprehensive solution built on logically related and consistent ideas, concepts, and attributes is what system architecture activities are aiming to achieve. The solution architecture has features, properties, and characteristics that, to the greatest extent possible, satisfy the problem or opportunity expressed by a set of system requirements (traceable to mission/business and stakeholder requirements) and life cycle concepts (e.g., operational, support).

These features, properties, and characteristics are implementable through technologies (e.g., mechanics, electronics, hydraulics, software, services, procedures, human activity).

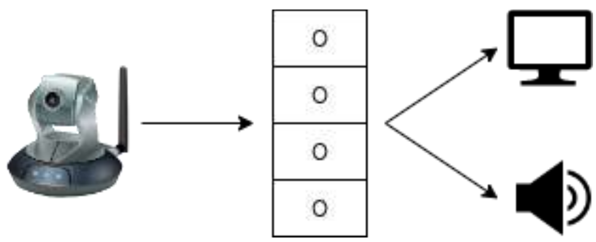

Fig.1. *Structural view of disabled to normal people*

Research and Reviews: Advancement in Robotics
Volume 5 Issue 3

This shows how handicapped people are structurally different from healthy people. The Raspberry Pi detects a person's sign captured by the camera and plays audio and text in response.

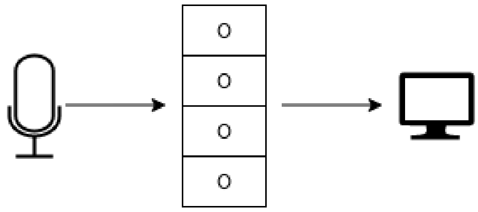

Fig.2. *Structural view of normal to disabled people*

The structural picture of people with impairments is shown in this diagram. The Raspberry Pi decodes audio and transforms it to text when someone speaks.

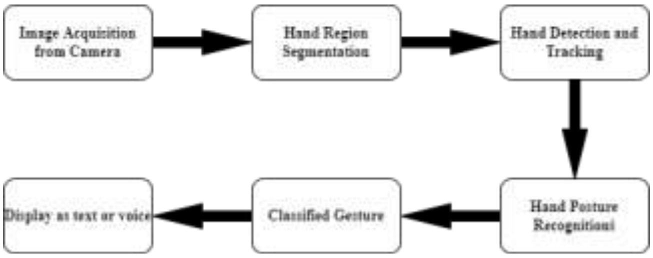

Fig.3. *Gesture Input from the camera*

This figure displays the camera's gesture input. Hand detection and tracking, hand posture recognition, and finally motion classification and display as text or voice take place after obtaining an image from a camera and segmenting the hand region.

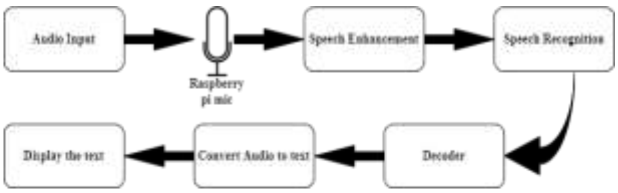

Fig.4. *Speech Input from microphone*

The graphic above shows the speech input from the microphone. After decoding the voice and turning the audio to text, the raspberry pi does speech enhancement and recognition before displaying the text.

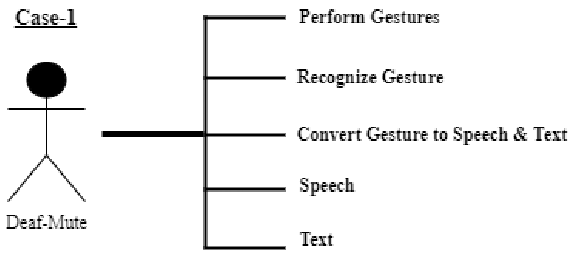

Fig.5. *Use Case Scenario for Deaf-Mute People*

A use case scenario for deaf-mute people is shown in the diagram above. The work flow of a deaf-mute is shown in this graphic.

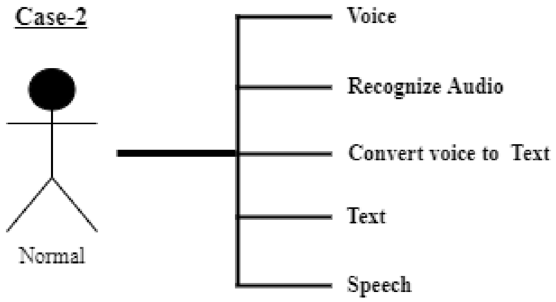

Fig.6*. Use Case Scenario for Normal People*

The diagram above depicts a use case scenario for deaf-mute people. This diagram depicts a deaf-work mute's flow.

The detailed design process determines the internal logic of each module mentioned in the high-level design. For each of the components, this step specifies more information and the algorithmic design. A lower level's additional components and subcomponents are also described in detail. Each portion will include a reference to this section or a detailed description of one of the system software components. Using flow charts for the most important modules, we provide much more detail on software modules and analyse the control flow in software using an organised diagram.

Makaton is a device with both hardware and software, and each piece of hardware houses a number of software programmes. This project suggests a camera-based assistance sign identification system that uses the Raspberry Pi, as well as processing some features including text and voice commands. Deaf and dump people typically use sign language to communicate.

RASPBERRY PI

It is a compact computer with a microprocessor inside. It can take many different shapes. We choose the Raspberry Pi 4 model due of its improved computing power, extra connectors (USB and HDMI), and more memory space. A quad-core ARM Cortex-A72 CPU with a 1.5 GHz clock speed powers the device. Raspbian is the operating system, and after the code has been run, it will read data from a webcam and send speech. A compact, inexpensive computer with complete capabilities is the Raspberry Pi. It is offered in a number of formats.

Fig.7. *Raspberry Pi*

The proposed system is put into practice using the Raspberry PI 3 model. It has an ARM cortex quad-core 64-bit processor. It has four USB ports and a 1GB internal memory. In addition, it includes built- in Bluetooth and Wi-Fi. The Raspberry Pi uses the ARM1176JZF-S core. All of the electronics required to run a computer can be placed on a single chip using a system on a chip (SoC). RAM is all compacted into one tidy package instead of having separate chips for the CPU, GPU, and USB controller. The Raspberry Pi needs an operating system to get going. In order to save money, the boot loaders are not stored in any on-board non-volatile memory on the Raspberry Pi. We analyse the acquired image with OpenCV in order to find the sign. for the boot loaders' sake. We use OpenCV to analyse the acquired image and find the sign.

The most recent model in the range of Raspberry Pi computers is the Raspberry Pi 4 Model B. It offers radically improved CPU speed, multimedia performance, memory, and connectivity over the Raspberry Pi 3 Model B+ of the previous iteration while preserving backward compatibility and consuming roughly the same amount of power. Performance-wise, the Raspberry Pi 4 Model B is comparable to that of entry-level x86 PCs.

IMPLEMENTATION

The implementation stage is the most important one in the development of any project since it yields the best solution to the issue at hand. During the implementation phase, the analysis from the analysis document and the creation of the anticipated phase are actually put into practice. Prior to the implementation of any software, important decisions must be made on the platform to be utilised, the language to be used, etc. Numerous factors, such as the actual environment in which the system functions, the required speed, additional implementation-specific properties, etc., frequently influence these decisions.

This chapter describes the programming language, environment, and other tools needed to complete the project and explains their importance to the system. These subjects are additionally explored in this chapter:

- Implementation Requirements.
- Guidelines for implementation.
- Implementation Procedure.

IMPLEMENTATION PROCEDURE

The implementation procedure is as follows:

DATABASE DESIGN AND CREATION

Signed languages use a variety of unusual encoding techniques. Employing various hand shapes, motions, wrist and palm orientations, using one or two hands, and facial expressions are a few of these methods. This dataset compares which of these grammatical building blocks are used in which sign languages.

One table is used to store information on hand shapes, and the other table is used to store additional data needed to make whole signs. This includes the hands' position, orientation, and shape in addition to the gestures that make up a sign. Both tables contain linguistic data as well as raw geometric data. The linguistic information provides independence from the geometry and access to lexical information for researchers, while the geometric information is used to create the three-dimensional graphical representation for a sign.

HAND DETECTION AND LOCALIZATION

A significant scientific challenge for improving communication with hearing-impaired people is computer recognition of sign language. In this paper, a quick and effective algorithm for counting the number of fingers spread out in an American Sign Language gesture representing an alphabet is introduced. Boundary tracing and fingertip detection are the basis for the idea of finger detection.

The system does not require the usage of any specific markers or input gloves on the hand, nor does it require the hand to be properly aligned to the camera. Index Terms: Finger detection, image processing, boundary tracing, sign language recognition, computer accessibility for the impaired.

- HAND RECOGNITION
- Image Capture: Scaling, Rotation, Translation, Lighting Conditions.
- Image Processing: Smoothing & Filtering.
- Feature Extraction: Hand Contour & Complex moment
- Classification: Evaluation & Comparison Hand Contour-based ANN and complex moments based-ANN

CONCLUSION

Sign language is the language of the deaf and hard of hearing. Without sign language, they could not function in the world. The development of human-machine interaction is assisted by gesture recognition technologies. The majority of scholars have previously concentrated on the identification of static hand gestures. There are some works that have been described for the recognition of dynamic hand motions. This uses a variety of acquisition strategies to recognise various sign languages. When using different data collection methods, these devices have a high rate of recognition, but due to their computational complexity and expense, they cannot be kept up in public spaces. Therefore, if you wish to avoid the significant issue, camera-based alternatives are preferable. Camera-based options are superior since they avoid the important problem. Using the Raspberry Pi camera, we quickly and simply processed the given data. In order to achieve better results and high accuracy, further research in the areas of segmentation techniques, feature extraction methods, and classification methods is needed to achieve the ultimate goal of human computer interaction in the field of sign language recognition for deaf and dumb people.

LIMITATIONS

Their shortcomings include being costly and difficult to use professionally. A multitude of classification strategies are employed by researchers. Researchers frequently develop their own theories based on time-tested methods to understand sign language more effectively.

FUTURE SCOPE

We can develop a model for ISL word and sentence level recognition. For this, a system that can detect changes in temporal space will be required. We can close the communication gap for persons who are deaf or hard of hearing by developing a comprehensive service. The proposed system can be developed and put into operation using a Raspberry Pi. We'll look for any motion-related clues.

REFERENCES

1. IEEE Journals in IEEE explore about: Sign Language Recognition Techniques- A Review 2020 IEEE International Conference for Innovation in Technology (INOCON) Bengaluru, India. Nov 6-8, 2020.
2. Mahesh Kumar N B, "Conversion of Sign Language into Text", in International Journal of Applied Engineering Research ISSN 09734562 Volume 13, Number 9 (2018) pp. 7154-7161.
3. Arsan, Taner and Ulgen, Oguz. (2015). Sign Language Converter.

International Journal of Computer Science and Engineering Survey. 6. 39-51. 10.5121/ijcses.2015.6403.

4. Rajesh George Rajan, M Judith Leo ISSN: 2277-3878, Volume- 7, Issue-6, March 2019.
5. K. Hogan, R. Stubbs, "Can't get Through 8 Barriers to Communication", Pelican Publishing Company, Grenta, LA, 2003.
6. Ms Kamal Preet Kour, Dr. Lini Mathew, Paper Id: IJTRS-V2-I7- 005, Volume 2 Issue VII, August 2017.
7. Sagar P.More, Prof. Abdul Sattar, "Hand gesture recognition system for dumb people", International Journal of Science and Research (IJSR)
8. Adithya, V. and Vinod, P.R. and Gopalakrishnan, Usha. (2013). Artificial neural network based method for Indian sign language recognition. 1080-1085. 10.1109/CICT.2013.6558259
9. Sunitha K. A, Anitha Saraswathi.P, Aarthi.M, Jayapriya. K, Lingam Sunny, "Deaf Mute Communication Interpreter- A Review", International Journal of Applied Engineering Research, Volume 11, pp 290-296, 2016
10. Mandeep Kaur Ahuja, Amardeep Singh, "Hand Gesture Recognition Using PCA", International Journal of Computer Science Engineering and Technology (IJCSET), Volume 5, Issue7, pp. 267-27, July 2015
11. X. Zhang, X. Chen, Y. Li, V. Lantz, K. Wang, J. Yang, "A framework for hand gesture recognition based on accelerometer and EMG sensors", IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans 41 (6) (2011) 1064– 1076.
12. J. Rekha, J. Bhattacharya, S. Majumder, "Shape, texture and local movement hand gesture features for Indian sign language recognition",3rd International Conference on Trendz in Information Sciences & Computing, 2011, pp. 30–35
13. Manikandan, K., Patidar, A., Walia, P. and Roy, A.B., 2018. Hand Gesture Detection and Conversion to Speech and Text,.arXiv preprint arXiv:1811.11997
14. Rahman, A. M., Ahsan, U., & Aktaruzzaman, M. (2011), Recognition static hand gestures of alphabet in ASL. Issn 2218- 5224 (Online), 2(1).
15. N. Pugeault, R. Bowden, "Spelling it out: Real-time ASL fingerspelling recognition", International Conference on Computer Vision Workshops, 2011, pp. 1114–1119
16. P.Vijayalakshmi and M. Aarthi, "Sign language to speech conversion," 2016 International "Conference on Recent Trends in Technology (ICRTIT), Chennai, 2016.
17. Chen, Zhi-Hua and Kim, Jung-Tae and Liang, Jianning and Zhang, Jing and Yuan, YuBo. (2014). Real-Time Hand Gesture Recognition Using Finger Segmentation. TheScientificWorldJournal. 2014.
18. Rajaganapathy, S. and Aravind, B. and Keerthana, B. and Sivagami, M. (2015). Conversation of Sign Language to Speech with Human Gestures. Procedia Computer Science. 50. 10.1016/j.pr