Real Time Bidirectional Communication Interpreter for Deaf Community

Om Panhale, Prasanna Phadtare, Shubham Patil, Sayali Chavan, Prof. Kirti Balsaraf
Department of Information Technology, Zeal College of Engineering and Research Pune

Abstract- Sign Language research field based on real time hand gestures called sign samples and recognition unit of computer. The sign language is the only way to used for deaf and dumb community communication platform. In this proposed system, we are working on the American Sign Language (ASL) dataset (A-Z) alphabet recognition followed by our word recognition dataset of Indian Sign Language (ISL). Sign data samples to be making our system more accurate with help of Convolutional Neural Network (CNN). Today, much research has been going on the field of sign language recognition but existing study has a limitation to develop trust over full communication interpreter. The purpose of this system is to represent a real time Hand Gesture Recognition (HGR) communication interpreter based on Indian Sign Language (ISL) with higher accuracy. Indian Sign Language (ISL) used by Deaf people's community in India, have acceptable, meaningful, essential and structural properties.

Index Terms- Convolutional Neural Network (CNN), Indian Sign Language (ISL), Hand Gesture Recognition (HGR), Deaf community (DC).

I. INTRODUCTION

There are so many languages in India used as officially and locally. Such large diversity country has more challenges to maintain uniqueness in language interpretation. In languages have its challenges when it used to communicate over different areas, societies and states. ISL is one of the living languages in India used by the Deaf community peoples. But as we seen there is not any standard language available till date. So we are working on different sign language dataset to invent Indian sign language as an interpreter.

We are going to implement two way communications for Sign language is used for the people who are deaf or who find it difficult to hear and also used by them who can hear but cannot physically speak. Our motive behind this implementation is to create entire language which involves functioning of hands, facial expressions and gesture of the entire body. The Sign language is not universal standard so we are making our contribution towards sign language development. Every nation has its self-developed sign language like American Sign Language work for alphabet recognizer. Each sign language has its own rule and semantic meanings.

The difficulty comes when deaf and dumb community want to communicate or trying to say something there is not any language for them. So it becomes essential to develop an automatic language interpreter to assist them for their fluent communication. They people want something more helpful which makes their communication universal and easy. The other one is based on computer vision based gesture recognition, which contains image processing techniques. So, this category faces more difficulty. our motive to develop this system based on real time signs.

This system captures hand gesture images of ISL with system camera for feature extraction. The analysis stage, pre-processing unit is used to the noise removal, grey scale conversion by using Gaussian filter, binary conversion of images done by using OTSU's method followed by feature extraction. In our system, CNN is going to used for future recognition in which we having the input unit of training data set of images. Next we have hidden unit which acts upon this training dataset to evaluate the output unit results train model. This entire CNN works by considering the factors namely matrix feature of images for drafting into a train model for real time sign recognition.

II. LITERATURE SURVEY

Sharmila Konwar [1] states that is used to model an automatic vision based American Sign Language detection structure and converting results in to text. The work introduced in this paper is meant to outlining the programming sight based American Sign Language recognition framework and interpretation to content. To distinguish the individual body skin shading from the picture, HSV shading model is utilized [1]. At that point edge recognition is connected to distinguish the hand shape from the picture. An arrangement of morphological activity is connected to get a refined yield for the gesture based communication acknowledgment This work is mainly focussed on the colour model and edge detection phenomenon. With the help of edge detection algorithm, the gestures are discovered successfully for the alphabets in American language. Some images are not discovered successfully due to geometric dissimiliarities, odd background and light states.

Yo-Jen Tu [2] presents a face and signal acknowledgment based on human-PC communication (HCI) framework utilizing a solitary camcorder. Not the same as the traditional specialized strategies among clients and machines, we unite head posture and hand motion to manage the hardware, so we can recognize the situation of the eyes and mouth, and utilize the facial focus to

assess the posture of the head. Two new techniques are explained in this paper: programmed signal territory division what's more introduction standardization of the hand signal. It isn't compulsory for the client to keep signals in upright position, the framework fragments and standardizes the signals consequently. They explore demonstrates this technique is extremely precise with motion acknowledgment rate of 93.6%. The client can control different gadgets, counting robots all the while through a remote system.

Angur M. Jarman [3] exhibits another calculation to distinguish Bengali Sign Language (BdSL) for perceiving 46 hand signals, including 9 motions for 11 vowels, 28 motions for 39 consonants and 9 motions for 9 numerals as indicated by the similitude of elocution. The picture was first re-sized and after that changed over to double configuration to edit the locale of enthusiasm by utilizing just best most, left-most and right-most white pixels. The places of the fingertips were found by applying a fingertip discoverer calculation. Eleven highlights were extricated from each picture to prepare a multilayered feed-forward neural system with a back-spread preparing computations. Separation between the centroid of the hand area and each fingertip was ascertained alongside the points between every fingertip and flat x pivot that crossed the centroid. A database of nearly 2300 pictures of Bengali signs happened to be developed to assess the viability of the proposed framework, where 70%, 15% and 15% pictures were utilized for preparing, testing, and approving, separately. Exploratory outcome demonstrated a normal of 88.69% exactness in perceiving BdSL which is particularly encouraging contrast with other existing techniques.

Javeria Farooq [4] states that hand motion acknowledgment was a characteristic and natural way to connect with the PC, since cooperation's with the PC can be expanded through multidimensional utilization of hand motions as contrast with other information techniques. Another methodology called "Arch of Perimeter" is given its application as a virtual mouse. The framework exhibited, utilizes just a webcam and calculations which are created utilizing PC vision, picture and the video handling tool stash of Matlab.

Guillaume Plouffe [5] examines the advancement of a whiz signal UI that tracks and perceives progressively hand signals in light of profundity information gathered by a Kinect sensor. The intrigue space relating to the hands is first portioned based on the suspicion that the hand of the client is the nearest protest in the scene to the camera. A novel calculation is proposed to move forward the checking time with a specific end goal to recognize the main pixel on the hand form inside this space. Beginning from this pixel, a directional scan calculation takes into account the recognizable proof of the whole hand form. The k-arch calculation is then utilized to find the fingertips over the form, and dynamic time twisting is used to choose motion competitors and furthermore to perceive motions by contrasting a watched motion and a progression of pre-recorded reference motions. The examination of results with cutting edge approaches demonstrates that the proposed framework beats a large portion of the answers for the static acknowledgment of sign digits and is comparable regarding execution for the static and dynamic acknowledgment of well-known signs and for the communication through signing letter set. The arrangement at the same time manages static and dynamic motions also similarly as with various hands inside the intrigue space. A normal acknowledgment rate of 92.4% is accomplished more than 55 static and dynamic signals. Two conceivable utilizations of this work are talked about furthermore, assessed: one for elucidation of sign digits and signals for friendlier human– machine cooperation and the other one for the normal control of a product interface.

Zafar Ahmed Ansari [6] states individuals with discourse inabilities convey in gesture based communication and accordingly experience difficulty in blending with the healthy. There is a requirement for a translation framework which could go about as a scaffold among them and the individuals who don't have the foggiest idea about their gesture based communication. An utilitarian unpretentious Indian gesture based communication acknowledgment framework was executed and tried on true information. A vocabulary of 140 images was gathered utilizing 18 subjects, totalling 5041 pictures. The vocabulary comprised for the most part of two-gave signs which were drawn from a wide collection of expressions of specialized and every day utilize starting points. The framework was executed utilizing Microsoft Kinect which empowers encompassing light conditions and question shading to have irrelevant impact on the effectiveness of the framework. The framework proposes a technique for a novel, minimal effort and simple to-utilize application, for Indian Sign Language acknowledgment, utilizing the Microsoft Kinect camera.

Keerthi S Warrier [7] states Hand Gesture Recognition System (HGRS) for detection of American Sign Language (ASL) has become compulsory and strong communication tool for particular user (i.e. hearing and speech impaired) to communicate with regular users via computer system. Different HGRS have been developed for finding of diversified sign languages using particular methods. There exist two main ways in the hand gesture analysis namely; vision-based and device-based way. In vision-based way the user does not need to wear any extra mechanism on hand. Else the system requires only camera(s), which are used to take the images of hand gesture symbol for communication between human and computers.

III. SYSTEM DESIGN

In this paper, we proposed a system which will overcome existing communication barrier by providing two way communications for deaf and dumb peoples. System input will be action of hand gestures and convert it into common words of communication after that converted text will be converted into voice. So that, normal people can understand it. Similarly, system will take input as voice and it will have converted into text and further convert it into sign language which is understandable by deaf people.

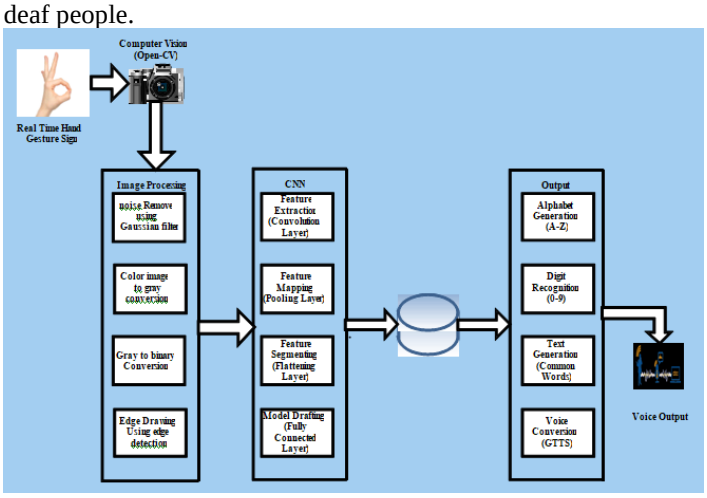

Fig.1: System Architecture

IV. SYSTEM MODULES

We are going to develop following modules:

1. Hand action recognition

Open-CV (Open Source Computer Vision) is a library of programming functions used for real time image processing with computer vision. In our implementation we use open compute vision for taking real time snap of hand gestures for further processing. After getting real time hand gestures image processing applied on it for removing noise from it.

2. Image processing

After getting real time image of hand gesture send for image processing module. In image processing, image gets converted in gray format by removing noise in it using Gaussian filter. After gray conversion image thresholding by setting RGB colour values to zero and preserving only black and white [0 and 1] values. Gray to binary conversion is done by using OTSU's method [1]. After getting black formatted image hand shape get extracted from image. The exact shape of hand will get by drawing edge using canny edge detection method.

3.Feature extraction

After getting exact shape of hand gestures features get extracted from it by using pixels' weight calculations. The image pixels get drafted in matrix by using weight gradient functions. Feature extraction done on all hand gesture dataset for training model creation and drafting. The train model creation done by using deep learning (CNN) algorithm.

4.Feature mapping & text generation

In real time image of hand gestures is going through image processing and subsequent phases of feature extraction. After getting image features these statistical features get matched with pre trained model and respective text generated. After text generation those text get converted into a voice.

5.Voice conversion

The text generated further gets converted into voice by using Google's text to speech library. After voice generation will be used for communication purpose deaf person to normal person.

6.Text to action conversion

For normal person to deaf communication normal person use their own language in the form of voice. The voice generated from normal people gets recognized by speech recognizer and this speech gets converted into a text. After text from normal person get semantically mapped with sign samples. The matched sign samples will show by using open-CV automatically. The sign images get easily understand by deaf and dumb community persons.

V. MATHEMATICAL MODEL

Process:

Let us consider S as a system for Sign Language Recognition System.

Mapping Diagram:

C1: sign inputs for Alphabet
C2: sign inputs for numbers
C3: sign inputs for words
R1: Result of sign recognition
R2: Result number recognition

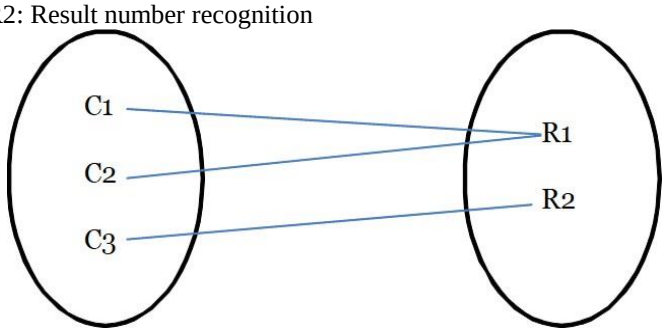

Fig.2: Mathematical Model

S={I,F,O}
INPUT:
F=F1,F2,F3...FN Function to execute result
I=C1,C2,C3... input of hand gestures
O=R1,R2,Rn...output

I=result access by User C1=sign recognition and voice conversion
C2= sign recognition and word conversion
C3= sign to number recognition
F: F1=sign recognition and voice conversion
F2= voice processing and sign conversion
O: R1= voice generation for recognized sign samples.
R2= sign suggestion for processed voice samples.

Above mathematical model is NP-Complete.
SPACE COMPLEXITY:
The space complexity depends on Presentation and visualization of discovered patterns. More the storage of data more is the space complexity.
TIME COMPLEXITY
Check No. of patterns available in the datasets= n
If (n>1) then retrieving of information can be time consuming. So the time complexity of this algorithm is O(nn).
θ = Failures and Success conditions.
Failures:

Huge database can lead to more time consumption to get the information.
Hardware failure.
Software failure.
Success:
High accuracy achieved by using American sign dataset.
User gets result very fast according to their needs

VI. EXISTING SYSTEM APPROACH

There are different approaches for image capturing is being used before. The captured images from either real time cam or static from dataset is used for further processing. Digital image processing is a field that analyses image processing methods. The image used in this is a static image form computer vision (webcam). Mathematically, the image is a formulation of light intensity on two-dimensional field. The image to be processed by a system or computer, an image should be presented statistically with numerical values. A digital image can be stated by a two-dimensional matrix f (m, n) consisting of M columns and N rows. The colour image processing [RGB], there are different models are like hue and saturation, value (HSV) model. This model is used with an object with a certain colour can be identified and to reduce the unwanted light intensity from the outside. Further Tests on images were performed using six kinds of colours, ie brown, yellow, green, blue, black and white.

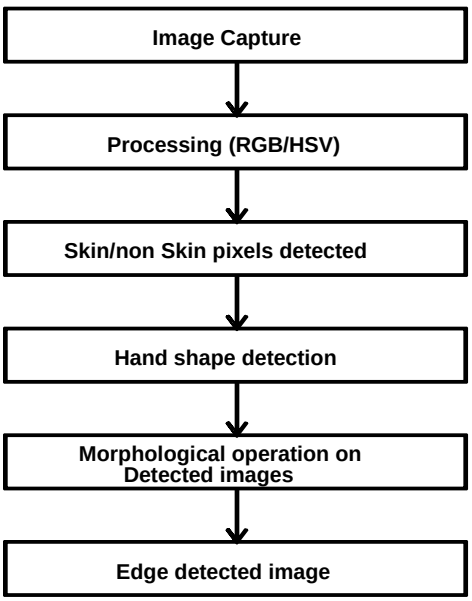

Fig.3: Existing System Architecture

A skin detector mainly used to transform a given pixel into an appropriate colour space. Then it is use a skin classifier to label the pixel. It correctly differentiates a skin or non-skin pixels. A skin classifier decides a decision boundary of the skin colour class in the color space based on a training database of skin-colours pixels. Hand shape detection done by using object recognition feature of edge drawing methods. The most primary morphological processes are used to adds pixels to the edges of hand shape in an image and eliminate pixels on hand shape object edges. The number of pixels added or eliminated from the hand shape in an image depends on the dimensions of real time sign image.

VII. PROPOSED SYSTEM APPROACH

Proposed system Sign language is the main language of the community who are deaf or facing difficulty of listening and also can be used by the ones who can hear but cannot speak. It is a hard but full language which contains functionality of hands, facial expressions and postures of the body. Sign language is not unique. Every country has its self developed sign language. Each sign language has its self defined rule of grammar, word orders and pronunciation. The problem arises when deaf and dumb community try to interact using this language with the people who are unknown of this language grammar. So it becomes essential to develop an automatic and interactive interpreter to grasp them.
So its mandatory to overcome these communication gap between the deaf community and normal persons. Two way communications system is providing for deaf and dumb peoples.

We are going to develop two-way communication systems by using machine learning and image processing techniques. The current real time application will work for real time assistance.

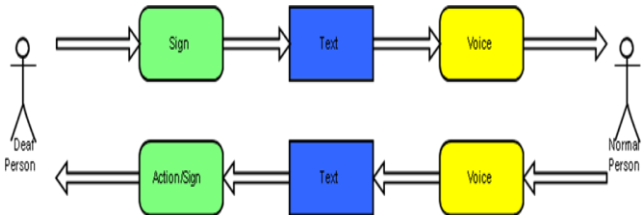

Fig.4: Proposed System Approach

VIII. METHODOLOGY USED

[1]. **OPEN-CV:**

Open-CV (Open Source Computer Vision) is a built in library of programming functions mostly aimed at real-time computer vision. In easy language it is library used for Image Processing. It is mostly used to do all the operation associated to Images.

[2]. **PYTHON:**

Python interface is being developed right now. There are many algorithms and many functions that write or bear those algorithms. Open-CV is developed in C++ and has a template interface that works faultless with STL containers.

[3]. **GTTS:**

It is a Python library and CLI tool to interface with Google Translates text-to-speech API. Writes spoken mp3 data to a file, a file-like object (byte string) for further audio manipulation, or studio.

IX. CONCLUSION

Thus the proposed system will overcome the problem. It will help to communicate between normal people and deaf people. System will be two-way communication system by using sign to text and voice to sign conversion phenomenon.

X. FUTURE WORK

Our future development will be extended for further improvement in detecting accuracy and also for motion detection of body for word recognition.

XI. REFERENCES

[1]. Sharmila Konwar, Sagarika Borah and Dr. T. Tuithung, **"An American sign language detection system using HSV color model and edge detection",** International Conference on Communication and Signal Processing, IEEE, April 3-5, 2014, India

[2]. Yo-Jen Tu, Chung-Chieh Kao, Huei-Yung Lin**,"Human Computer Interaction Using Face and Gesture Recognition**", Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific, IEEE, Kaohsiung.

[3]. Angur M. Jarman, Samiul Arshad, Nashid Alam and Mohammed J.Islam, **"An automated Bengali sign language recognition based on finger tip finder Algorithm",** International journal of Electronics & Informatics, 2015.

[4]. Javeria Farooq and Muhaddisa Barat Ali, **"Real time hand gesture recognition for computer interaction",** International Conference on Robotics and Emerging Allied Technologies in Engineering (ICREATE), 22-24 April ,2014.

[5]. Guillaume Plouffe and Ana-Maria Cretu, **"Static and dynamic hand gesture recognition system in depth data using dynamic time warping"** IEEE Transactions on Instrumentation and Measurement (Volume: 65, Issue: 2, Feb. 2016)

[6]. ZAFAR AHMED ANSARI, **"Nearest neighbour classification of Indian sign language gestures using kinect camera"** , February 201 , pp. 161–182

[7]. Keerthi S Warrier, "**Software based sign language converter"** , April 2016 IEEE.