Real Time Identification of American Sign Language for Deaf and Dumb Community

Snehal Pawar[1], Pragati Salunke[2], Arati Mhasavade[3], Aishwaraya Bhutkar[4], K. R. Pathak[5*]
[1,2,3,4]*Student,* [5]*Assistant Professor*
Department of Computer Science,
Rajgad Dnyanpeeth Technical Campus Dhangawadi,
Pune, India.
Corresponding Author
E-mail Id:- krpathak121213@gmail.com

ABSTRACT

The only way for deaf and dumb for communication is based on sign language which involves hand gestures. In this system, we are working on the American Sign Language (ASL) dataset (A-Z), (0-9) and word alphabet identification escort by our word identification dataset of Indian Sign Language (ISL). Sign data samples to be making our system more faultless, error free, and unambiguous with help of Convolutional Neural Network (CNN). Today, much research has been going on the field of sign language recognition but existing study failed to develop trust full communication interpreter. The motivation of this system is to serve a real-time two-way communication translator based on Indian Sign Language (ISL) with higher precision, efficiency, and accuracy. Indian Sign Language (ISL) used by Deaf-mute people's community in India, does have adequate, delightful, acceptable, meaningful essential and structural properties.

Keywords:-*Artificial Neural Network, Indian Sign Language, Hand Gesture Recognition, Deaf community.*

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 transmit or communicate over different zone, civilization and situation. Indian Sign Language (ISL) is one of the daily bread languages in India used by the Deaf-mute 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 hard of hearing and also used by them who can hear but cannot physically speak. Our motive behind this implementation is to create complete language which involves movement of hands, facial expressions and gesture of the body. The Sign language is not universal standard so we are making our contribution towards sign language development. Every country has its own native sign language like American Sign Language work for alphabet recognizer. Each sign language has its own rule and semantic meanings. The problem comes when deaf and dumb people want to communicate or trying to say something there is not any language for them. So it becomes necessary to develop an automatic language interpreter to assist them for their fluent communication. They people want something more helpful which makes there communication universal and easy. Another one is based

on computer vision based gesture recognition, which involves image processing techniques. Consequently, this category faces more complexity. 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 analyzing phase, 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, Convolutional Neural Network (CNN) is going to use 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. The working with real time sign language we know that the dataset need to be large and rich in processed features.

LITERATURE SURVEY
Sharmila Konwar et. al. [1] states that is used to design an automatic vision based American Sign Language detection system and converting results in to text. The tasks introduced in this wrapper is meant to plan a programmed vision based American Sign Language perception structure and presentation to content. To immortalize the human skin shading from the picture, HSV shading mode is utilized. At that point edge detection is connected to immortalize the hand shape from the image. An justification on study of shape activity is connected to get a common output for the sign based conversation salute This task is mainly entrap on the color mode and edge diagnosis phenomenon. Boundary diagnosis (edge detection) algorithm the hand action is discovering successfully for

the alphabets in American language. Some action is not detected successfully due to geometric variations, odd background and light conditions.

Yo-Jen Tu et. al. [2] presents a face and signal cognizance based human-PC communication (HCI) structure utilizing a reclusive camcorder. Not the same as the old specialized blueprint among clients and machines, we amalgamate head posture and hand motion to sway the hardware. We can recognize the situation of the eyes and mouth, and utilize the facial focus to assess the posture of the head. Two new methods are show in this wrapper: programmed signal sector division what's more, introduction standardization of the hand action (signals). It isn't mandatory for the client to keep action in upright position, the scheme tatter and standardizes the signals accordingly. They explore manifest this method is tremendously precise with gesture recognition rate of 93.6%. The client can control different apparatus, including robots all the while through a remote system.

Angur M. Jarman et. al. established Bengali Sign Language (BdSL) for purpose of getting 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 pronunciation. In this first size of picture was changed and after that changed done to double configuration for making changes to the locale of enthusiasm by using best most, left-most and right-most white pixels. Fingertip discoverer calculation was applied on it for detecting fingertips. Eleven highlights were extracted from every picture for preparing a multilayered feed-forward neural system with a back-spread preparing calculation. Separation between the centroid of the hand area and each fingertip was discovered with the points between each fingertip and flat x pivot that

crossed the centroid. A database of 2300 pictures of Bengali signs was developed to assess the numerous 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 DBs which is particularly encouraging contrast with other priory techniques.

Javeria Farooq et. al. states that hand motion is a characteristic and natural way to interact with the PC, since association with the PC can be broaden through multidimensional usage of hand motions as contrast with other information techniques. The reason behind this paper is to find out three unique strategies for HGR (hand signal acknowledgment) utilizing fingertips location. Another technique called "Arch of Perimeter" is given its application as a virtual mouse. The framework exhibited, uses just a webcam and calculations which are based on utilizing PC vision, picture and the video handling tool stash of Matlab.

Guillaume Plouffe et. al. stated the of a whiz signal User interface that perceives progressively hand signals in light of depth reason information gathered by a Kinect sensor. The interested space relating to the hands is first portioned based on the suspicion that the hand of the client should be near to camera. A novel calculation is aim to move forward the checking time with a specific destined aim to recognize the main pixel on the hand form inside this space. From first time from this pixel, a directional scan calculation takes into account the identifiable proof of the whole hand form. The k-arch calculation is then established to discover 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 development 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 related execution for the static and dynamic acknowledgment of well-known signs and for the communication with the help of signing letter set. The arrangement at the same instant manages static and dynamic motions also similarly as with number of 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 et. al. proposed individuals with essay inabilities convey in gesture based communication and as per experienced difficulty in blending with the healthy. There is a need for a changing framework which could go about as a scaffold among them and the individuals who not have the hazy idea about their gesture based communication. A utilitarian unpretentious Indian gesture based communication acknowledgment framework was established and tried on real information. A vocabulary of 140 images was collected using 18 subjects, total 5041 pictures. The vocabulary consists 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 using Microsoft Kinect which empowers inefficient impact on the well-deserved of the framework. The framework contains 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 et. al. states Hand Gesture Recognition System (HGRS)is strong and useful communication media

for establishing communication with deaf and dumb people. Lots of HGRS have been developed for discovering diversified sign languages using appropriate techniques. There are two main methods in the hand gesture analysis namely; vision-based and device-based approach. In vision-based approach the user does not require to wear any extraneous mechanism on hand. Instead the system requires only camera(s), which are aim to capture the images of hand gesture symbol for communication between human and computers system.

SYSTEM DESIGN
We are going to overcome existing communication barrier by providing two way communications for deaf and dumb peoples. We take input as action of hand gestures and convert it into common words of communication after getting this in text, convert it into voice. After getting voice, normal people can understand it. Similarly, normal people can speak in voice; our system will convert it into text and further convert it into actions which are simply understandable by deaf people.
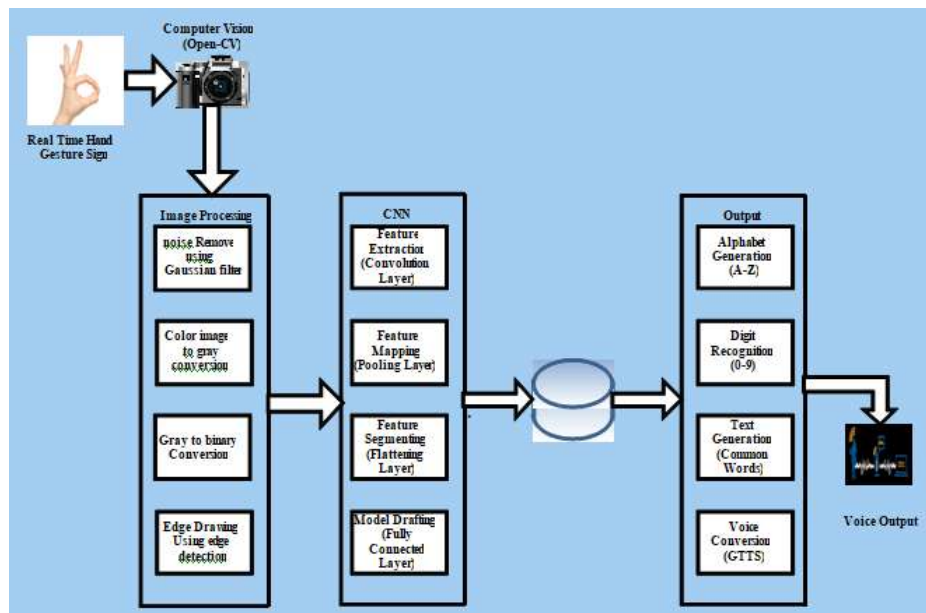

Fig.1:-*System Architecture*

We are going to develop following modules:
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 are going to 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.

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. 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.

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.

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.

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.

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.

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.
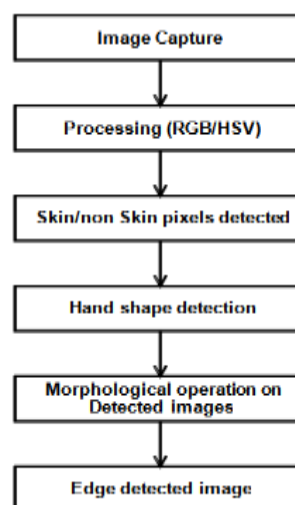

Fig.2:-Existing System Architecture

A skin detector mainly used to transforms 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 basic morphological operations are used to add pixels to the boundaries of hand shape in an image and removes pixels on hand shape object boundaries. The number of pixels added or removed from the hand shape in an image depends on the size and shape of real time sign image. After studying multiple papers we conclude that those all work is unable to create communication interpreter while they are suffering or focusing only image processing and object recognition.

PROPOSED SYSTEM APPROACH
In this system Sign language is the key language of the people who are deaf or who are not able to hear and also used by them who are able hear but unable to speak. It is a difficult but complete language which contains motion of hands, facial expressions and poses of the body.

Sign language is not a global language. Every country has its own local sign language. Each sign language has its own grammar rules, word structure and sounding. The problem exists when deaf and dumb people try to communicate using with normal people who are unable to understand this language. So it becomes important to develop an automatic and reciprocal transcriber to understand them.

So it's mandatory to overcome these communication gaps between the deaf community and normal persons. Two way communications system is providing for deaf and dumb peoples. We take input as action of hand gestures and convert it into common words of communication after getting this in text, convert it into voice. After getting voice, normal people can understand it. Similarly, normal people can speak in form of voice; our system will convert it into text and further convert it into actions which are simply understandable by deaf people.

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.
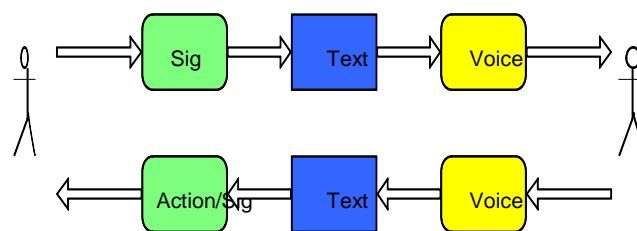

Fig.3:-Proposed System Approach

METHODOLOGY USED
Open-CV
Open-CV (Open Source Computer Vision) is an open source computer vision and machine learning software library. It is a cross platform library used to develop real time computer vision applications. It mainly focuses on image processing, video capture and analysis including features like face detection.

Python
Python interface is being actively developed right now. There are many algorithms and many functions that compose or support those algorithms. Open-CV is written natively in C++ and has a template interface that works

Image Processing
Read and write images and detection of images for extracting its features. Detection of shapes like Circle, rectangle etc. in an image, Detection of coin in images. Image processing works with text recognition in images. e.g. Reading Number Plates, Modifying image quality and colors.
e.g. Cam-Scanner.

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.

CONCLUSION
In this system, we are going to implement two way communication interface between normal people and deaf and dumb people to make communication easier. So we are going to propose hand gesture recognition systems based on American Sign Language dataset with our Indian sign contribution along with some common words using deep learning approach. System will be two way communication system by performing sign to text and voice conversion and voice to sign conversion.

FUTURE WORK
Our future work will be for further improvement in sign identification accuracy and also for motion detection of hand gesture word identification

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 fingertip 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.2016.65(2).
6. ZAFAR AHMED ANSARI, "*Nearest neighbour classification of Indian sign language gestures using kinect camera*", February 2016.41(2):161–182p
7. Keerthi S Warrier, "*Software based sign language converter*", April 2016 IEEE.