**Smart Aid For Speech And Hearing Impaired Community Using CNN**

Dr. BVANSS Prabhakar Rao prabhakarrao@vit.ac.in

Lavanya Shivani lavanya.shivani2015@vit.ac.in

Manisha Chaudhary manisha.chaudhary2015@vit.ac.in

Akash

akash2015@vit.ac.in

***Abstract*-**Deaf and dumb people uses gesture based communication to communicate with the world using hand gestures. Sign language is the only single way of communication for deaf and dumb people. But common people face difficulty in understanding the gesture language therefore often these physically challenged people has too keep the translator with them to communicate with the world. The progression in implanted framework can give a space to plan and build up an interpreter framework to change over the communication via gestures into discourse. These days implanted framework has turned into a critical pattern in all applications. Our prototype can be used for the welfare of the society . Firstly for the travelling purpose , dumb people face problems while expressing . So they can use our translator which gets converted into text thus helping the pedestrians or the common people. This can be used the either way too that is conversation amongst different groups of deaf and dumb. Different group refers to deaf and dumb people knowing different sign languages as (ASL,ISL,BSL etc.). Secondly , this can be used for the shopping purpose when the mute people can't reply to the

owner or the deaf people can't hear to the directions shown through audio. Also , for the meeting purposes where one of the co-worker wants to share his/her views with (HR) but can't speak so here our prototype comes into the picture . Now let us take the situation where the HR who is dumb wants to convey some regards to the one who is deaf, to which he can use our prototype which could be helpful enough to make a successful meeting . So we would be using Convolutional Neural Network (CNN) which would implement our prototype .Thus reducing the pre-processing and enhancing the overall recognition process. The work exhibited fundamentally lessens the correspondence gap amongst imbecilic and standard individuals and intends to encourage dumb individual's way of life.

**Keywords : CNN , ASL ,BSL, ISL , Precision , Recall , F1-Score**

**I.INTRODUCTION**

Motion of anyone part like face, hand is a type of Gesture. Here for gesture acknowledgment we are utilizing CNN. Gesture acknowledgment empowers PC to comprehend human activities and furthermore goes about as an translator among PC and human. This could give potential to human to connect normally with the PCs with no physical contact of the mechanical gadgets. Gestures are performed by deaf and dumb to perform gesture based communication. This people group utilized communication via gestures for their correspondence when broadcasting sound is outlandish, or composing and composing is troublesome, yet there is the vision probability. Around then communication through signing is the way to trade data between individuals. Ordinarily communication through signing is utilized by everybody when they would prefer not to talk, however this is the main method for correspondence for hard of hearing and dumb. Communication through signing is likewise serving a similar importance as spoken language does. This is utilized by hard of hearing and dumb community everywhere throughout the world yet in their local structure like ISL, ASL. Sign language can be performed by utilizing Hand signal either by one hand or two hands. It is of two kind Isolated gesture based communication and constant communication via gestures. Separated communication through signing comprises of single gesture having single word while nonstop ISL or Continuous Sign language is a grouping of gestures that produce a significant sentence. In this report we performed ASL gestures recognition system.

**1.1 Objective**

To create an effective communication tool for the people who are not able to speak or hear anything. A User Friendly, Cost effective system which reduces communication gap between deaf and dumb with ordinary person. This framework offers voice to voiceless i.e. voice is given to the individual who can't talk. Imbecilic/quiet individuals utilize gesture based communication for correspondence reason. Communication through signing utilizes signals rather than sound to pass on data. This dialect incorporates consolidating hand shapes, hand developments, outward appearances to express person's considerations. Some of them are thevariety of the hand motion appearance, scaled, pivoted adaptation of picture and the picture handling speed as it includes numerous scientific estimations. Generally, there are diverse methodologies presented to perceive hand signal.

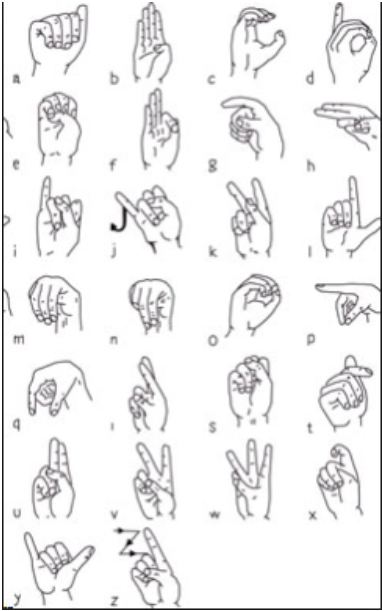
**1.2 Background**

Gesture recognition and sign language recognition has been a well researched topic for American Sign Language(ASL), however few research works have been distributed with respect to Sign Language(ISL). One of the methodologies included key point discovery of Image utilizing SIFT and afterward coordinating the keypoint of another picture with the keypoints of standard pictures per letters in order in a database to arrange the new picture with the mark of one with the nearest coordinate . Another determined the eigen vectors of covariance framework determined from the vector portrayal of picture and utilized euclidean separation of new picture eigen vector with those in preparing informational collection to characterize new picture . However , rather than utilizing top of the line innovation like gloves or then again kinect, we plan to take care of this issue utilizing CNN.

**1.3 Commonly Used Sign Languages**

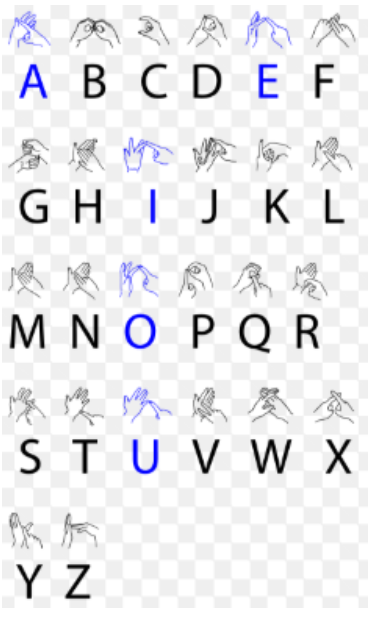
On moving across the globe we came to know that there is no universal sign language . usage of sign language varies from community to community so we cannot stick to one particular language . Let us deal few of them one by one. Namely ASL(American Sign Language ), ISL(Indian Sign Language ) , BSL (British Sign language) etc.

ASL - Other than North America, lingos of ASL and ASL-based creoles are utilized in numerous nations around the globe, including quite a bit of West Africa and parts of Southeast Asia. ASL is likewise generally learned as a second language.



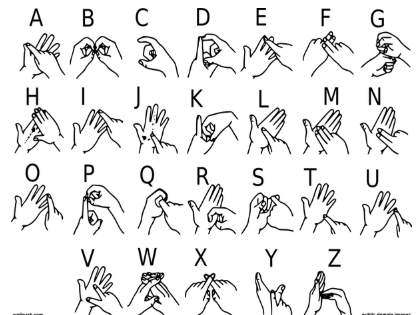
**Fig.1. Alphabets Of ASL**

**ISL - Indian Sign Language .**



**Fig. 2 . Alphabets Of ISL**

**BSL -** BSL uses two handed alphabet. Mainly used in [Scotland](https://www.google.com/search?q=Scotland&stick=H4sIAAAAAAAAAONgVuLQz9U3MCs0LFrEyhGcnF-Sk5iXAgCFrvSvFwAAAA&sa=X&ved=2ahUKEwiehPPIsa7hAhVMfH0KHUiuD1UQmxMoATAiegQIDBAH), [England](https://www.google.com/search?q=England&stick=H4sIAAAAAAAAAONgVuLQz9U3MMqqMFzEyu6al56TmJcCALMnJfIWAAAA&sa=X&ved=2ahUKEwiehPPIsa7hAhVMfH0KHUiuD1UQmxMoAjAiegQIDBAI), [European Union](https://www.google.com/search?q=European+Union&stick=H4sIAAAAAAAAAONgVuLSz9U3iDcria-yWMTK51palF-QmpinEJqXmZ8HAKuzoYEfAAAA&sa=X&ved=2ahUKEwiehPPIsa7hAhVMfH0KHUiuD1UQmxMoAzAiegQIDBAJ) and Wales.



**Fig. 3. Alphabets Of BSL**

**II. LITERATURE SURVEY**

In the ongoing years, there has been gigantic research on the hand sign language gesture recognition. The technology for gesture acknowledgment is given beneath.

### 2.1 Sensor Based

The proposed model will be consisting of combination of hardware and software. Hardware part will include flex sensors on each finger, microcontroller, power supply, and android phone and Bluetooth module. Software part will include programming for android phone application. Hardware part will be consisting of flex sensors to take input from different gestures through gloves, microcontroller to convert input analogue data to digital data and for further processing, power supply to provide voltages to specific units, and finally Bluetooth module to send the data from controller to android mobile. We propose an algorithm to describe the operation of the system.

ALGORITHM

Below is the algorithm of the proposed system

Step 1. Start

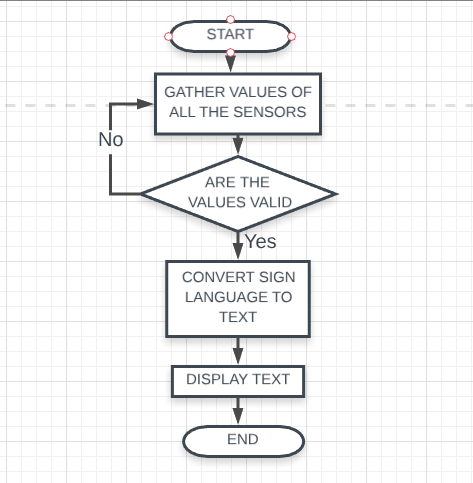
Step 2. Gather Values from the sensors.

Step 3. Are values forming any meaning? If yes then sends conversion of sign language to text data to Bluetooth. Android app gathers data from Bluetooth abd displays it.

Step 4. Android app gathers data from Bluetooth and displays it.

Step 5. Stop

#### 



#### Fig. 4. Flow Model (Sensor Based)

**2.2 Using Image Processing**

**ALGORITHM**

Step1.Reading Image from the camera and applying pre-processing techniques like gamma correction, blurring.

Step2. Hand Segmentation using background subtraction algorithm.

Step3.Hand detection using thresholding and dilation.

Step4. Finding Contours of hand for getting shape of hand.

Step5.Finding contour Area, Convex Hull, hull area Solidity.

Step6. Also find the angle between fingers and the aspect ratio of hand.

Step7. Finding the defects of hand using convex hull.

Step 8.Finally classifying using solidity,aspect ratio, convex defects and angles.

Step 9.If image(sign)==database image, then ON the LED and speak the meaning of that sign.

Repeat Step 9

END

**2.3 Vision Based**

In vision­ based techniques PC camera is the info gadget for watching the data of hands or fingers. The Vision Based techniques require just a camera, along these lines understanding a characteristic cooperation among people and PCs without the utilization of any additional gadgets. These frameworks will in general supplement biological vision by depicting artificial vision frameworks that are executed in programming as well as equipment. This represents a difficult issue as these frameworks should be foundation invariant, lighting insensitive, individual and camera autonomous to accomplish constant execution. Also, such frameworks must be improved to meet the prerequisites, including accuracy and robustness. Vision based examination, depends on the way human beings perceive information about their surroundings , yet it is most difficult to implement in a way that would be quite satisfying. A few unique methodologies have been tried up until now.

1. One is to build a three ­dimensional model of the human hand. The model is coordinated to the image of the hand by at least one cameras, and parameters  relating to palm orientation  and joint angles are evaluated. These parameters are then used to perform gesture classification.

2. Second one to capture the image using a camera then extract some feature and those features are used as input in a classification algorithm for the classification.

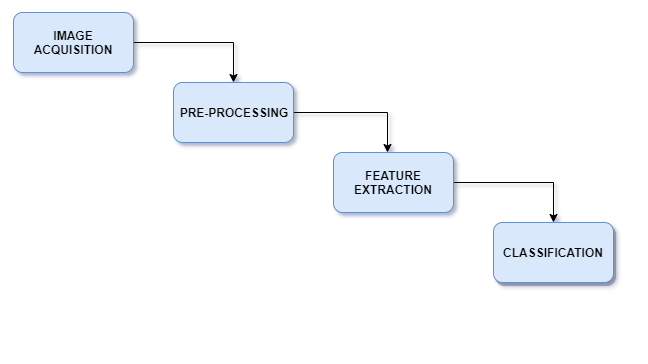
****

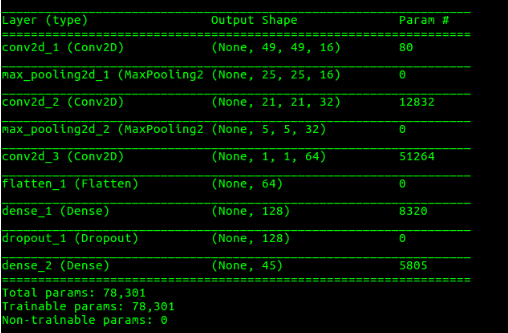
Fig. 6. Vision Based

**III.PROPOSED SYSTEM**

The proposed system is designed to which successfully able to recognize the hand patterns when the images are taken as input as well when the image is captured using live motion capture.

**3.1 Using CNN**

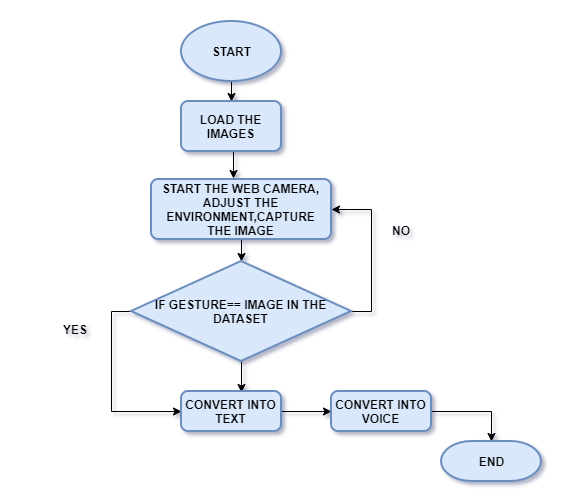
Convolutional neural networks aim to use spatial information between the pixels of an image. Therefore, they are based on discrete convolution. CNNs is the most popular neural network model being used for image classification problem. The big idea behind CNNs is that a local understanding of an image is good enough. The practical benefit is that having fewer parameters greatly improves the time it takes to learn as well as reduces the amount of data required to train the model. Instead of a fully connected network of weights from each pixel, a CNN has just enough weights to look at a small patch of the image. It’s like reading a book by using a magnifying glass; eventually, you read the whole page, but you look at only a small patch of the page at any given time. On analyzing the above approaches we further come to a conclusion that in Sensor-based approach normally underwriter requires to wear gloves in hand, additionally a few sensors ,etc that are utilized as sign signal to demonstrate the hand stance or motion, while visionbased methodology endorser does not require to wear anything regular hand utilized as sign flag to display hand stance or signal. We are keen on vision-based methods to perceive hand pose, which are increasingly normal method for correspondence with incapacitated individuals and robots. further on proceeding with vision based we came into a dilemma of two approaches firstly image classification algorithms secondly CNN. Since image classification algorithms requires preprocessing which is the reason why handling the data is more time taking and have several constraints . Due to above loop holes we come to a conclusion that CNN would give the more accurate results.



**Fig 7. Model Summary**

**2.2 Gesture Recognition**

The Final step is recognition(matching), Keypoints between two images are matched by identifying their nearest neighbours. But in some cases, the second closest-match may be very near to the first. It may happen due to noise or some other reasons. In that case, ratio of closest-distance to second-closest distance is taken. If it is greater than 0.8, they are rejected. It eliminaters around 90% of false matches while discards only 5% correct matches, as per the paper.



**Fig. 8. Flow Model (CNN)**

**3.3 Deploying the system on cloud**

This model relies on remote servers for processing logic that is accessed through a web browser with a continual internet connection. As Cloud Computing has grown in popularity, several different models and deployment strategies have emerged to help meet specific needs of different users. Each type of cloud service, and deployment method, provides you with different levels of control, flexibility, and management. Understanding the differences between Infrastructure as a Service, Platform as a Service, and Software as a Service, as well as what deployment strategies you can use, can help you decide what set of services is right for your needs.

The System uses IaaS (Infrastructure as a service) model for its deployment. Google cloud platform(GCP) is used for installing of the system.

Step1: create an account on Google Cloud Platform.

Step2: Create a Virtual machine over GCP. Install the required operating system. (windows/Linux)

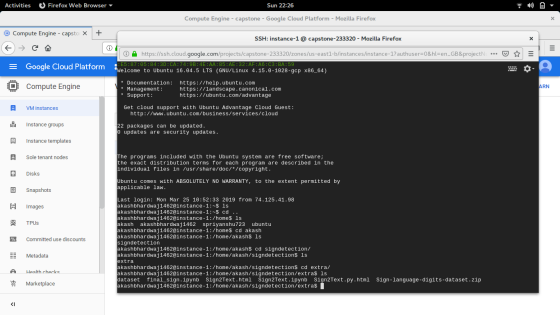
Step3:Run the program over cloud. The image/video will be taken through camera which will be sent over to cloud system for processing and the matched output will be sent over to the user back in form of audio and led display.

Benefits of using cloud :

1.Easy Deployment

2.Fast response to business needs. Cloud applications can be updated, tested and deployed quickly, providing enterprises with fast time to market and agility. This speed can lead to culture shifts in business operations.

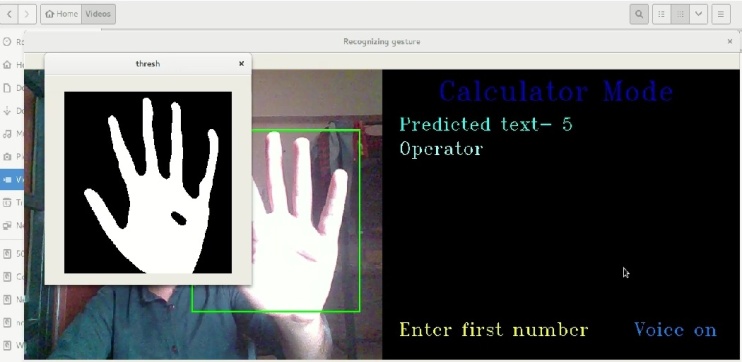
3.Instant scalabilityAs demand rises or falls, available capacity can be adjusted.



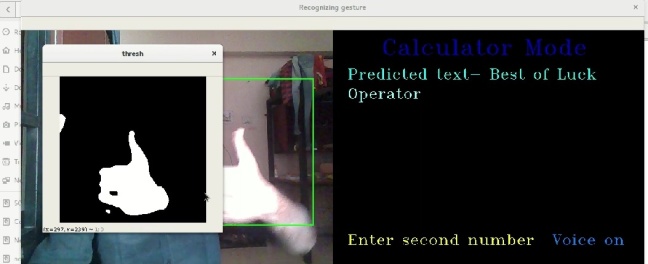
**Fig. 9. Cloud Enviornment**

**IV.RESULT**

The Gesture was Recognized successfully and converted to text and audio.

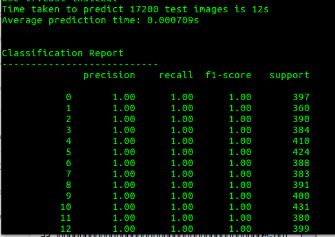
****

**Fig.10. Output 1**



**Fig.11. Output 2**

Precision score for the images recognized.

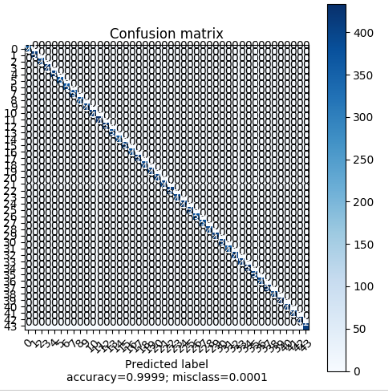


**Fig.12. Precision/Recall**

**Validation and Loss Visualization**

**Fig.13. Accuracy And Loss Graph**

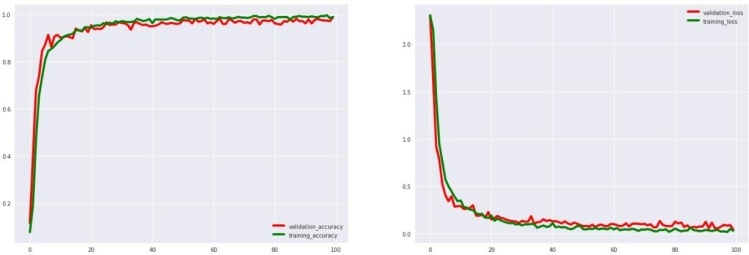
**Confusion Matrix**:Confusion matrix, also known as an error matrix, is a specific table layout that allows visualization of the performance of an algorithm.



**Fig. 14 . Confusion Matrix**

**VI. CONCLUSION AND FUTURE WORK**

The propose system was successfully implemented. The system was able to help deaf and dumb people in their daily communications. The deployment of system to cloud provided a fast and efficient method to solve the compatibility and performance issues which were raised while installing the software to every customer side.

The above project can be taken for further work by creating web browser based API’s model which will make the deployment of system very effective.

Many applications can be developed using sign language inputs such as developing a python calculator, using input in gaming industry where hand gestures can be used to control the game modules.

**VII. REFERENCES**

[1] Intelligent Sign Language Recognition Using Image Processing Sawant Pramada1, Deshpande Saylee 2, Nale Pranita3, Nerkar Samiksha4 Mrs.Archana S. Vaidya “Sapat College of Engineering, Management Studies and Research, Nashik (M.S.)”

[2]https://www.researchgate.net/publication/220791651\_A\_Survey\_on\_Sign\_Language\_Recognition

[3] Sign Language Recognition Using Image Processing Kamal Preet Kour, Dr. Lini Mathew Electrical & Panjab University, Punjab, India.

[4] Glove Based hand gesture Recognition- a product of Cyber Glove systems, USA

[5] Sign Language Recognition Using Image Processing Kamal Preet Kour, Dr. Lini Mathew Electrical & Panjab University, Punjab, India

[6]https://opencvpythontutroals.readthedocs.io/en/latest/py\_tutorials/py\_feature2d/py\_sift\_intro/py\_sift\_intro.html

[7] https://aws.amazon.com/types-of-cloud-computing/

[8] https://searchcloudapplications.techtarget.com/definition/cloud-application

[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, Issue 7, pp. 267-27, July 2015.

[11] Sagar P.More, Prof. Abdul Sattar, “Hand gesture recognition system for dumb people”,International Journal of Science and Research (IJSR), Volume 3, Issue 2, April 2015.

[12] Pratibha Pandey, Vinay Jain, “Hand Gesture Recognition for Sign Language Recognition: A Review”, International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 3, March 2015 .

[13] W. K. Chung, W. Xinyu, and Y. Xu. 2008 “A realtime hand gesture recognition based on Haar wavelet representation”, in Proceedings of the IEEE International Conference on Robotics and Biomimetics, Washington, DC, USA, 2008, pp.336-341.

[14] M. P. Paulraj, S. Yaacob, H. Desa, and W. Majid. 2009 “Gesture recognition system for Kod Tangan Bahasa Melayu (KTBM) using neural network”, in 5th International Colloquium on Signal Processing and Its Applications, pp. 19- 22.

[15] Witkin, A.P. 1983. Scale-space filtering. In International Joint Conference on Artificial Intelligence, Karlsruhe, Germany, pp. 1019-1022.

[16] Lowe, D.G. 1999. Object recognition from local scale-invariant features. In International Conference on Computer Vision, Corfu, Greece, pp. 1150-115.