Hand Gesture Recognition for Deaf and Mute

Praveen Kumar, Tushar Sharma, Seema Rawat, Saksham Bhagat

Abstract*: Communication is the fundamental channel between individuals to speak with each other. In the world of communication gestures and signals is a great deal and a lot of research work has been done in the course of past decades. With the end goal to enhance the recognition rate of systems, numerous specialists have conveyed strategies, for example, HMM, Artificial Neural Networks, and Kinect stage. Effective algorithms for segmentation, classification, pattern matching, and recognition have evolved. Gesture-based communication is generally utilized by people with hearing disabilities to speak with one another helpfully utilizing hand motions. The system uses image processing technology and neural networking for the capturing and conversion of gestures. A number of python packages are used to process and generate the results. The application uses laptop webcam for capturing gestures and recognize gestures shown by the user. The application uses TensorFlow and Keras to generate the model for datasets. The gestures shown by the user are compared with stored gestures and the corresponding output is generated along with speech output. The application thus eliminates the communication barrier between hearing impaired-mute and normal people.

***Keywords*: Gestures, Recognition System, Image Processing, Neural Network, TensorFlow.**

I. INTRODUCTION

CNN is a class of neural networks that has the most common application in the field of analyzing visual images. CNN comprises of "neurons" which fulfill the purpose of taking in multiple inputs, communicating the set of inputs to all the functions that make use of them and produce a single output. Image processing is one of the most rapidly developing areas of technology in various applications for example, biometrics, secure communication, data translation and processing, biomedical, pattern recognition and many more [6]. Image processing and recognition is a boon for disabled people who cannot communicate properly with others and this technology can also be useful when there is a difference of language between speaker and receiver as it can act as a translator. Gesture-based communications are the most the natural and characteristic form of language that could be gone back to a time as the approach of human development, when the first sign of language showed up ever. The main challenges that this special person facing is the communication gap between -a special person and a normal person. Deaf and Dumb are individuals dependably discover challenges to speak with typical individuals. This gigantic test makes them uneasy and they feel segregated in society. Due to miss correspondence, Deaf and Dumb individuals feel not to convey and consequently they never ready to express their emotions [7].

II. LITERATURE SURVEY

Zimmerman et al. [1] worked on a hand gesture recognition device i.e. a hand to machine cross-point device responsible for providing real-time gesture, spot and orientation information required. This works with the help of a glove and various different technologies. This application using a glove is utilized in combination with a computer that hosts and drives a #d model of the hands giving the glove bearer the power to operate computer-simulated objects as if they are real and also providing a platform for a visual programming language**. Kurakin et al**. [2] also worked on a similar project proposing a real-time system for dynamic hand recognition with proper use on sensors. They worked in the area of depth-sensing for the understanding of human activity and behavior. With great work comes challenges where they employed a new technique hand orientation detection. The mentioned system was evaluated on a dataset of 12 dynamic ASL gestures. Fang et al. [3] tried to overcome the orthodox way of computer-human interaction i.e. keyboard, mouse and bringing in light the vision-based hand interaction using gesture recognition. The method requires a specific gesture to initiate the hand recognition followed by tracking, then the hand is divided into segments using color cues and at the last stage, the space-scale feature is used in gesture recognition to prevent the limitation of aspect ratio from occurring. This method was finally applied to the navigation of image browsing giving desired results. **Pickering et al**. [4] in this field emphasizes a research study on hand gesture recognition technologies for human and vehicle interaction, describing a primary and secondary task of driving together with HMI trends. hence this project offers a great combination of hand gesture recognition which may even offer some added safety benefits for secondary controls. **Freeman and Roth** [5] present a method to recognize hand gestures using a technique by engaging histograms of local positioning. This method is faster and simpler to compute and offers robustness as mentioned by the authors. They have also implemented it in distinguishing small vocabularies by 10 different hand gestures and also look up to implement the same in various future projects such as video games.

III. METHODOLOGY

Technologies Used: -

1. OpenCV 2. Pandas 3. NumPy 4. TensorFlow
5. Keras 6. H5py 7. Pyttsx38. CNN
9. Grayscale image processing 10. Spyder IDE

Revised Manuscript Received on December 05, 2019.

Dr. Praveen Kumar, Student: Associate Professor, Computer Science & Engg, Amity University Uttar Pradesh , Noida, India
Dr. Seema Rawat, Student: Associate Professor, Information Technology, Amity University Uttar Pradesh , Noida, India
Tushar Sharma, Student, Computer Science & Engg, Amity University Uttar Pradesh , Noida, India
Saksham Bhagat, Student,Computer Science & Engg, Amity University Uttar Pradesh , Noida, India

1) OpenCV- It has optimized libraries which include computer vision and machine learning algorithms. These functionalities are then used to detect motion, recognize faces, identify human actions, track movements, identify objects, finding results for a searched image from the database, etc.

2) Pandas- Pandas allow one to load data from different sources into Python and then use Python code to analyze those data and produce results which can be in the form of tables, text and also visualization, with the help of visualization.

3) NumPy- Python stores and reads images using arrays of numbers. And here NumPy is used. So, as we represent images with a list which is not very efficient for big images and a list of these big images occupy a lot of memory and therefore slowing down operations on them. So, this is solved by NumPy, a Python library that provides a multi-dimensional array object. It is a fundamental package for computing in Python.

4) TensorFlow- TensorFlow is an open-source programming library for superior numerical calculation. Its adaptable engineering permits simple sending of calculation over an assortment of stages (CPU s, GPU s), and from desktops to groups of servers to portable and edge gadgets.

5) Keras- Keras is a Python deep learning library written in python and consists of high-level neural network API. Due to its characteristics such as modularity, extensibility and friendliness allow fast and easy prototyping. It runs seamlessly between CPU and GPU s. The main purpose of this library is to enhance fast experimentation.

6) H5py- It gives you a chance to store huge measures of numerical information, and effortlessly control that information from NumPy. For instance, you can cut into multi-terabyte datasets put away on disk, as though they were genuine NumPy exhibits. A huge number of datasets can be put away in a solitary record, sorted and labeled anyway you need.

7) Pyttsx3- An OFFLINE Python Text to Speech library (TTS) which works for both python3 and python2.This library exceptionally useful particularly if you don't need any delay in the speech created and don't have any desire to depend just on the web for TTS change.

8) CNN- A CNN can be created using three simple steps: Convolution, Activation, Pooling.

Repetition of steps 1-3 helps to create a better CNN by having a greater number of hidden layers. The main application of CNN is in the field of image classification. It can deal with a large dataset say up to 100,000 images, learn the important features that help in making correct predictions.

9) Grayscale image processing- It produces high accuracy when searching for faces because when we have a BGR image with a profusion of features OpenCV will not be working with 100% accuracy leading to missing out of faces or even features. However, using a Grayscale image as an alternative to BGR increases accuracy.

10) Spyder IDE- Spyder is an open-source cross-platform IDE for data science. It offers a unique combination of debugging, analysis and advance editing functionalities of a complete development tool with the deep examination and beautiful conception capabilities of a scientific package.

IV. PROPOSED MODEL

The work on the project has been done using Machine Learning and Deep Learning concepts, making use of standard python libraries and packages to compile the neural network. To capture and store the gestures, first we need to segment the background color from the color of our hand using skin detection and segmentation methods After successfully segmenting the color, we need to draw and store the contour the hand gesture which defines the features of particular hand shape. So, in general, we have our proposed work is consist of 2 parts:

A preprocessing step: -

In this step, the segmented image is processed by using the probabilistic value of the threshold. Now, as we are using OTSU Binarization in our project, so the threshold value is calculated of the value of the mean of the histogram of the image to determine the skin likelihood [3]. Through that, we determine which pixel has skin likelihood and which don't and as a result, we have a threshold image showing white as skin color and black as the background color.

Extraction of features and characteristics: -

The contour produced by hand will act as characteristic for the gestures and define its features. The handshape will go about as the element of the motion. The extraction part of picture analysis tries to recognize natural features or highlights of objects found inside a picture. These qualities are then utilized to portray the object, or characteristic of the item, before the consequent undertaking of classification. [8]. The figure below represents the working flowchart of our application that defines each of the components of our application.

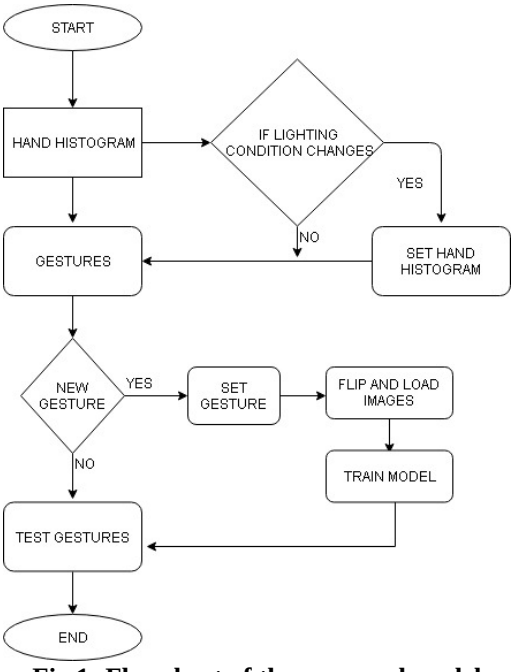

Fig.1: Flowchart of the proposed model

When the application starts for the first time, it requires a histogram to be stored. If the user is changing its lighting conditions, he/she have to again run the hand histogram script. The images are then captured[14] and go for the processing to remove noise and enhance the features of the image.

After successfully setting up the histogram, the user can create its own gestures or add new gestures to old ones.
When gestures are stored, then they are flipped so that gestures can be made or recognize by either hand and loaded with images and labels respectively. In the next step, the user has a choice to go for which trained model either go for TensorFlow or go for Keras. Finally, after training the data, the user tests the gestures with the trained model and get the desired output in the form of text and speech.

V. RESULTS

The model is consisting of a number of libraries that generate some outputs and acts as input for next set of instructions.

1) **Setting hand histogram-** The application starts with a pre-sampling step to separate colors of the background and your hand. It collects colors using squares on the screen. These squares calculate the threshold to get binary data. We need to put the hand in squares and making sure the hand covers all the squares.

.

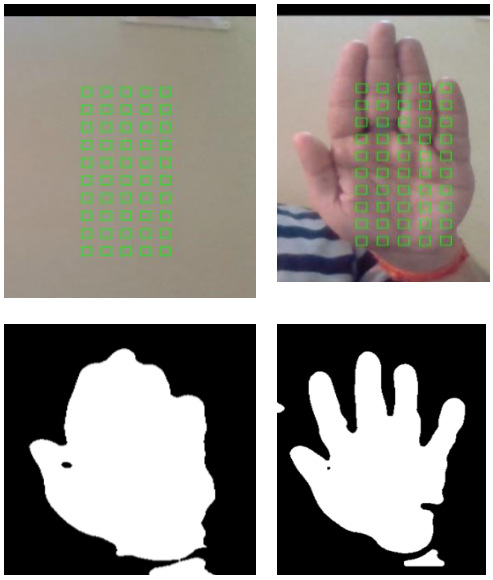

Fig.2: Setting up the hand histogram

2) **Creating hand gestures-** When running the script, it will ask for gesture number and gesture name to store gestures in folder for each gesture. After entering both the input values, two screens will appear.

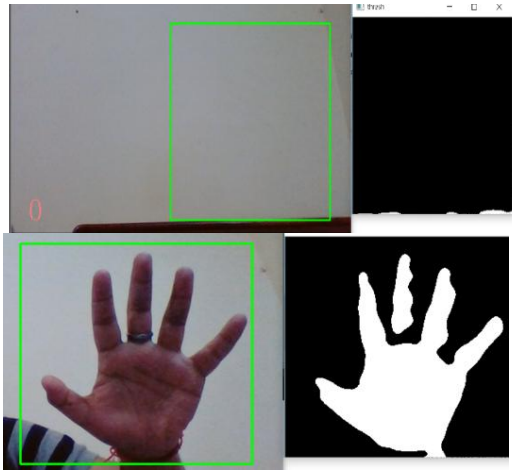

Fig.3: Creating hand gestures

3) **Flipping and Load Images-** The pictures were flipped using flip_images.py. This script flips every image along the vertical axis. Hence each gesture has 2400 images stored inside the gesture/folder. The above script appends the images with their labels and determines the length of train images and train labels.

4) **Training the model-** The script is used to train the model for the stored gestures, in the first section train and test images of a gesture[12][13] is loaded into the program and then these images are sent to the cnn_model () function. Basic CNN functions are implemented such as convolutional step, pooling and flattering step. In this last step, optimizer optimizes the model using several stored gestures and create a Keras model as cnn_model_keras.h5 in the root folder of the application.

5) **Recognizing gestures and running text to speech engine-** This is the script which uses gesture to communicate between deaf and mute persons. We need to perform gestures inside the green box and the program will determine the appropriate selection from its trained datasets and produce output in the right hand side of the window [10]. These scripts also using the screen to speech packages so every text obtained from gestures is converted to sound and outputted to system speakers. You can also input words and the same letters by specifying the gestures continuously.

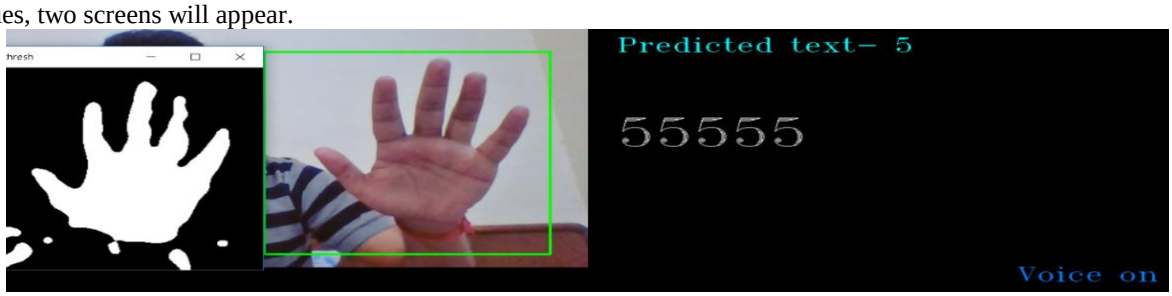

Fig.4: Testing the gestures

VI. ANALYSIS OF RESULTS

All input images/ gestures are stored using a laptop webcam. The captured image is then converted into binary data using the NumPy package and processed by using the H5py package. Now using the algorithm and functions from OpenCV packages, the application can give out the matched hand gestures from its trained model. It also produces the audio output using the device speaker indicating output sign gesture.

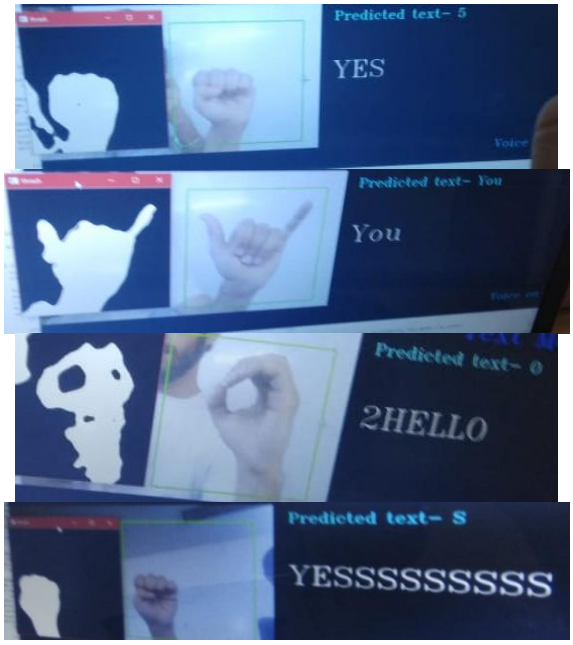

Fig.5: Outputs

Table 1: Test Results with Webcam

Gesture	The recognition rate for 5 trials using a webcam (approx.)
1	70
2	70
3	60
4	80
5	90
C	80
G	70
L	90
O	80
Y	90

VII. CONCLUSION

The application is tested and developed successfully with a laptop webcam. The application is useful for deaf and mute people with the system having a webcam connected to it. Some gestures have a recognition rate between 80-90% while some of them have a lower rate (not less than 40) but overall the application has approx. 80% accuracy with a laptop webcam. The performance of the application is depending on the algorithm used to develop the model and detecting sign images, also the performance is greatly varying due to lighting conditions and noise background which can be reduced by using the high-performance webcam. The gestures can be stored according to user requirements or can use stored gestures provided with the application.

One of the real difficulties experienced today is giving equivalent facilities to the differently-abled and impeded. While there are uncommon necessities and arrangements around the globe, there's as yet immense hole for development to bring the change in the lives of such individuals. In communication via gestures, each motion as of now has relegated meaning, and solid principles of setting and language structure might be connected to make recognition tractable. Gesture Recognition will dispense with a lot of manual assignments and will make life a lot simpler [9].

VIII. FUTURE SCOPE

This will certainly impact the future with many positive applications. Some may include:

- Medical: Gesture recognition can be integrated with advance robotics to be kept in homes and hospitals to detect and prevent life threatening situations like heart attacks. These can also be paired with public CCTV cameras to predict possible accidents and call for an ambulance on time.
- Additional Computer Interface: Gesture recognition, coordinated with voice acknowledgment, face recognition, lip movement catching and iris recognition together can be utilized to make perceptual UI (PUI), an extraordinary method to associate with PC frameworks which will propel its convenience and advancement.
- Entertainment: In the world of entertainment videogames play a vital role and as the world is advancing so for the need of a perfect control for the same. As the world moves towards the virtual reality and with huge scope to improve, the need for gesture recognition to be used as controls also increases acting as primary and secondary inputs and making tasks a lot easier and fun to do.
- Hand gestures can be used for analyzing and annotating video sequences of technical talks.

REFERENCES

1. Lanier, J., Blanchard, C., A hand gesture interface device, SIGCHI Bulletin (Vol. 18, No. 4, pp. 189-192). ACM.
2. Zhang, A real time system for dynamic hand gesture recognition with a depth sensor. EUSIPCO (Vol. 2, No. 5, p. 6), 2012, August.
3. Fang, Y., Wang, K., Cheng, J., & Lu, H. (2007, July). A real-time hand gesture recognition method. In Multimedia and Expo, 2007 IEEE International Conference on (pp. 995-998). IEEE.
4. Burnham, K. J., Hand gesture recognition technologies and applications for human vehicle interaction, 3rd IET Conference on (pp. 1-15) 2007.
5. W. T, Orientation histograms for hand gesture recognition. In IEEE FG (Vol. 12, pp. 296-301) 1995.
6. Nitesh S. Soni ; M. S. Nagmode, Hand gesture recognition & classification for deaf & dumb, ICICT,26-27 Aug. 2016.
7. Manisha U. Kakde , Amit M. Rawate, Hand Gesture Recognition System for Deaf and Dumb People, IJESC, Volume 6 Issue No. 7, 2016.
8. Bradski, Gary, and Adrian Kaehler. Learning OpenCV: Computer vision with the OpenCV library. " O'Reilly Media, Inc.", 2008.
9. Manisha U. Kakde ," Hand Gesture Recognition System for Deaf and Dumb People Using Principal Component Analysis ", IJESC, July 2016, Volume 6 Issue No. 7.
10. Aswathy M , Heera Narayanan , Surya Rajan , Uthara P M , Jeena Jacob," Hand Gesture Recognition and Speech Conversion for Deaf and Dumb using Feature Extraction", IJAREEIE, Vol. 6, Issue 3, March 2017.
11. Image Processing in OpenCV," https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_imgproc/py_thresholding/py_thresholding.html".
12. Jaswal, Kalpana, Praveen Kumar, and Seema Rawat. "Design and development of a prototype application for intrusion detection using data mining." In 2015 4th international conference on reliability, infocom technologies and optimization (ICRITO)(trends and future directions), pp. 1-6. IEEE, 2015.
13. Rawat, Seema, Praveen Kumar, and Geetika Jain. "Implementation of the principle of jamming for Hulk Gripper remotely controlled by Raspberry Pi." In Proceedings of Fifth International Conference on Soft Computing for Problem Solving, pp. 199-208. Springer, Singapore, 2016.
14. Kumar, Praveen, Sanchita Kadambari, and Seema Rawat. "Prefetching web pages for improving user access latency using integrated Web Usage Mining." In 2015 Communication, Control and Intelligent Systems (CCIS), pp. 401-405. IEEE, 2015.

AUTHORS ' S PROFILE

Dr. Praveen Kumar, is working as Associate Professor at Amity School of Engineering & Technology. He is M.Tech in Computer Science & Engineering. He has a number of international and national publications to his credit. He is a lifetime member of IETE, ACM, and IET. His primary research area includes Big
Data Analytics, Cloud Computing and Data mining.

Mr. Tushar Sharma, is a student at Amity School of Engineering Technology, Amity University, Uttar Pradesh. His area of interest includes machine learning, big data, convolution neural network and IOT

Dr.Seema Rawat, is working as Assistant Professor at Amity School of Engineering & Technology. She is PhD and M.Tech in Computer Science and B.Tech in Information Technology. She has a number of international and national publications to her credit. She is a member of IEEE, IACSIT and IAENG. Her primary research area includes Cloud Computing, Data mining and Artifi cial Intelligence.

Mr. Saksham Bhagat, is a student at Amity School of Engineering Technology, Amity University, Uttar Pradesh.Leveraging technology for a better tomorrow is his area of interest. His area of research includes convolutional neural network and python deep learning.