International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

# Sign Language Recognition

**E.Padmalatha, S.Sailekya, R.Ravinder Reddy, Ch.Anil Krishna ,K.Divyarsha**

***Abstract:*** ***There are nearly 15 million people around the world who have difficulty in speaking or communicating. Their only way of communication is through sign language. Hand gesture is one of the methods used in sign language for non-verbal communication. It is most commonly used by deaf & dumb people who have hearing or speech problems to communicate among themselves or with normal people. There are many recognized sign language standards that have been defined such as ASL(American Sign Language), IPSL(Indo Pakistan Sign Language), etc., which define what sign means what. ASL is the most widely used sign language by the deaf and dumb community. The deaf and dumb use sign language to communicate among themselves with the knowledge of the standard sign language. But they can't communicate with the rest of the world as most of the people are unaware of the existence and the usage of the sign language. This method aims to remove this communication barrier between the disabled and the rest of the world by recognizing and translating the hand gestures and convert it into speech.***

***Key words: sign language, disabled, nonverbal communication***

## I. INTRODUCTION

The hearing disabled and mute individuals can't blend with the social world as a result of their disabilities. Involuntarily they are treated differently by the society. They cannot perform well in many areas of interaction. For example, education environment for person with disabilities is not similar as the rest of the people, disabled people do not have any special tools to buy commodities, they have hard time to find work, and much more. It creates a gap between person with and without the disabilities. This gap is ever increasing day by day. There are more than 10 million deaf adults and deaf children in the world using American Sign Language (ASL) as a way of communication. In spite of such large numbers, very little efforts have been put to bridge the gap.

The sign language is a very important way of communication for deaf-dumb people. In sign language each gesture has a specific meaning. So therefore complex meanings can be explain by the help of combination of various basic elements. Sign language is a gesture based language for communication of deaf and dumb people.

Manuscript published on 30 September 2019
* Correspondence Author
**Dr.E.Padmalatha*,** M.Tech(CSE),Ph.D(CSE), Assistant Professor, CBIT Hyderabad, India
**S.Sailekya,** B.E(CSE) Student BVRIT, Hyderabad, India
**Dr.R.Ravinder Reddy,** M.Tech(CSE),Ph.D(CSE) Associate Professor, CBIT Hyderabad, India
**Ch.Anil Krishna**, B.E(CSE),CBIT Hyderabad, India
**K.Divyarsha**, B.E(CSE),CBIT Hyderabad, India



It is basically a non-verbal language which is usually used by deaf and dumb people to communicate more effectively with each other or normal people. Sign language contains special rules and grammars for expressing effectively.

Basically there are two main sign language recognition approaches image-based and sensor based. But lots of research is going on image based approaches only because of advantage of not need to wear complex devices like hand gloves, helmet etc. like in sensor based approach. Gesture recognition is gaining importance in many applications areas such as human interface, communication, multimedia and security. Typically Sign recognition is related as image understanding. It contains two phases: sign detection and sign recognition. Sign recognition is recognizing a certain shape that differentiates the object from the remaining shapes. Especially in the cases when no alternative communication is available. The technical point of view characteristic features of sign language communication are: its social direction and meaning; technical and technological convenience and easy to use. The system will use a webcam for the capturing images and pre-processing of the signs will be done by using OpenCv library. On having the input sequence of images captured through web-cam here the model uses some image preprocessing steps for removal of background noise and employs slope distance based algorithm i.e. Fingertip Detection by which generates a ratio with a help of which a template of the captured image is generated.

## II. METHODOLOGY

The main objective of the proposed system is eliminating the communication barrier and eases the process of communication between able and disabled people. Along with recognizing the gesture, we also give a speech output of the interpretation. There are nearly 15 million people around the world who have difficulty in speaking or communicating. Their only way of communication is through sign language. Hand gesture is one of the methods used in sign language for non-verbal communication. It is most commonly used by deaf & dumb people who have hearing or speech problems to communicate among themselves or with normal people.

The deaf and dumb use sign language to communicate among themselves with the knowledge of the standard sign language. But they can't communicate with the rest of the world as most of the people are unaware of the existence and the usage of the sign language. Our project aims to recognize and translate the hand gestures and convert it into speech.
Having observed the number of deaf and dumb people[1] who have difficulty in communicating and the existing systems that don't provide a complete end to end system to eliminate the issue, we decided to develop this solution.

- A human mediator between the two people.
- Sign language recognition using sensor gloves.

## III. IMPLEMENTATION OF THE PROPOSED SYSTEM

**Identifying hand color**

The main preprocessing step involves hand segmentation which includes identifying just the hand from the entire image and

1. Create a region of interest on the screen where the user must put his hand.
2. Calculate the histogram of the region of interest and save it. These values hold the color of hand.
3. Threshold values for HSV are calculated based on the histogram obtained. All the values above this are made white and below this are made black.
4. Whenever a new image is fed to the system, all the regions of the new image matching the previous histogram are made white and rest black. This way, the hand is segmented.

**To create gestures**

The next step is to create ASL hand gestures. This algorithm is used to already create a dataset of 26 alphabets, 10 numbers and 11 phrases which users can use directly. The user can create new gestures and add to the dataset with this.

5. Take input of gesture number and name. Save it in the database.
6. Get the skin color details from set_hand_hist.py output and identify hand from the entire image.
7. Apply necessary filters to the image and start capturing images until 1200 images are captured in various angles with slight modifications.
8. Save the images and run flip_.images.py to store flip all the images.

**Training the model**

After capturing the images of the new gesture[3],the images are flipped to recognize left hand gestures also. Then the CNN model is trained using these images.

9. Run load_images.py to get two folders, train_images with 5/6th of total and test_images with 1/6th of total images.
0. Create a Convolutional Neural Network model using Keras module of python.
11. Choose a sequential model with Convolutional, Pooling and Dense layers.
12. Save the model and the results can be observed in model.png.

**Identifying the gesture**

1. Load the keras model [4] and get the histogram values.
2. After processing the hand image, pass it to mode.predict() to get the probabilities.
3. The result is a list of probabilities of which the highest is the classifier's output.
4. Once probability is greater than 70% and the gesture is consistent for more than 15 frames, the model goes to database and fetch the corresponding text of that gesture.
5. The text is then converted to speech using pyttsx3 module.

The proposed system main aim is to recognize the American Sign Language and convert it into text and then into speech.

The first step is to isolate the hand of the user from the entire image. For this, in the initial stages of development, set_hand_hist.py is used to identify the user's hand's skin color. All the objects of this color are made to white and rest to black, thereby isolating the hand from background. As this method fetched subpar results in low light conditions, we used a pink colored glove to avoid the recognition of face and other body parts that have same color as the hand. This helped in obtaining sharper boundaries while isolating the hand.

Using this process, the dataset of various gestures has been created which include alphabets, numbers and few phrases. Each gesture has been recorded in various positions with a count of 1200 images and the flipped for better classification giving 2400 images per gesture. These are then given to the keras based CNN[10] model. There are 3 convolution layers, 2 pooling layers, 1 flatten layer, 2 dense layers and 1 dropout layer. 5/6th of the images(94000) are used for training and remaining 1/6th of the images(18800) for the testing of the data. 20 epochs have been run over the entire training set, and observed consistent increase in accuracy until 99.4% on the training data. The total time for training took over 80 minutes. The confusion matrix obtained after the training gives the accuracies with which each of the gesture was classified during the training stage.

## IV. RESULTS AND DISCUSSION

In this system, a dataset containing all the gestures are present. Each gesture folder consists of 2400 images which is used for training and testing the model. There are 47 gestures but more can be added by the users.

The above Figure 2 shows the 47 Folders belonging to 47 different gestures with 26 letters, 10 numbers and a few. Each folder consists of 2400 greyscale pictures of every gesture at different positions.


**Figure 2 Folders Present in the Dataset**

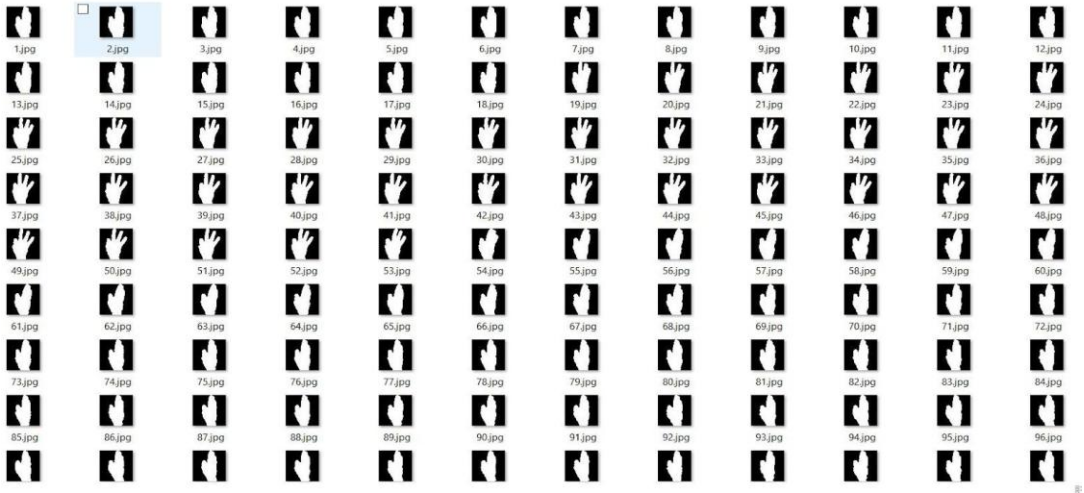

**Figure 3 Sample of Files present in folder 5**

Figure 3 shows the files present in folder 5. It shows all the different pictures of the gesture at various positions. We have 47 folders with 2400 images each for each gesture.

*Screenshot of capturing hand color*

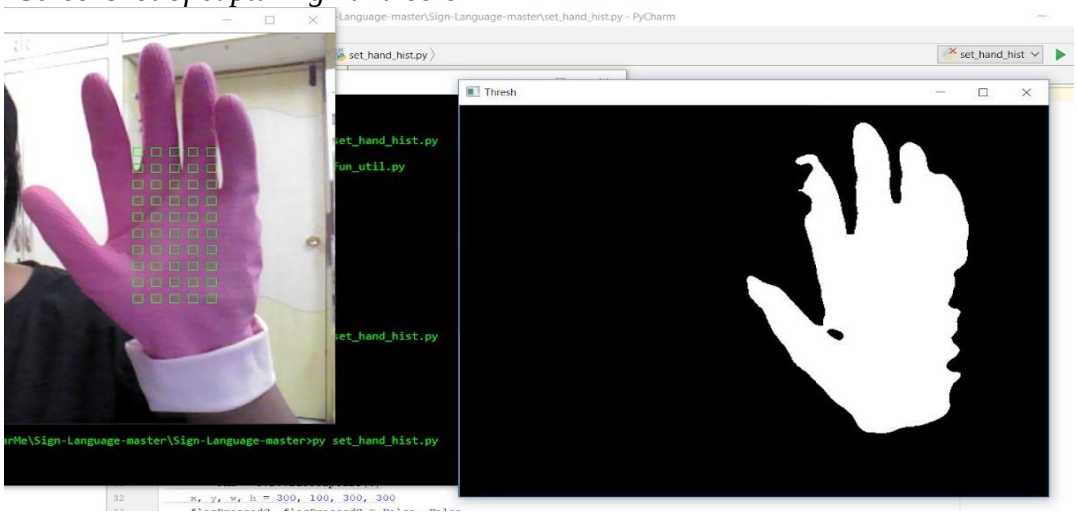


**Figure 4 Capturing hand color**

Figure 4 shows the screenshot of the project capturing the hand color. The model identifies the colors of those 50 boxes and generates color parameters. which are then used to whiten portion of those color and blacken the other portions.

*Screenshot of creating gestures*


**Figure 5 Executing create_gesture file.**

The above Figure 5 shows the execution of create_gestures.py file. It will ask the user to enter the gesture number and the title of the gesture. If the gesture number already exists, it'll ask the user to change/keep the record.

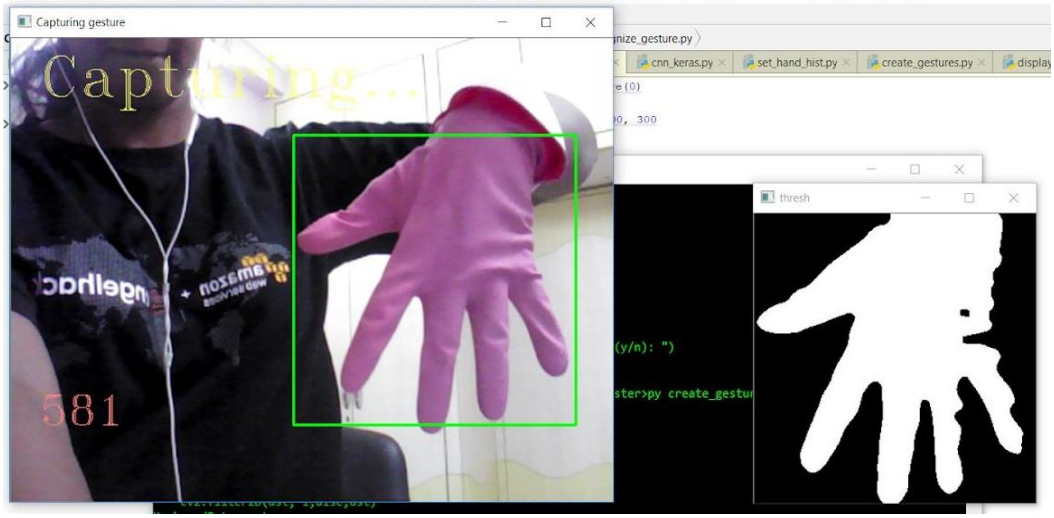


**Figure 6 Capturing a new gesture.**

The above Figure 6, shows the capturing of gestures. The number on the left bottom indicates the number of images captured. The model captures 1200 images of the new gesture and saves it in a file.

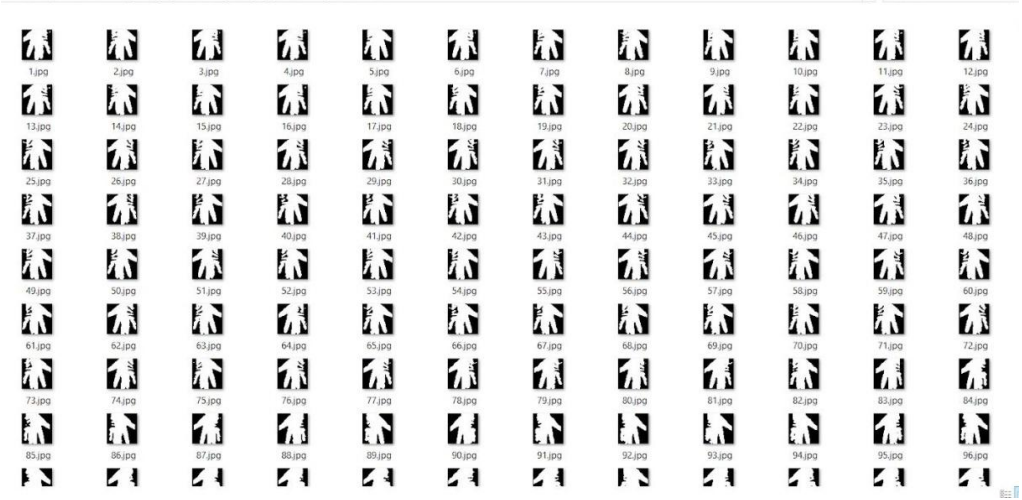

**Figure 7 Images of new gestures saved in a file**

Figure 7 is the file with the number 46 in which all the newly created gestures are saved. There are 2400 images as the 1200 images captured are flipped in order to recognize the same gesture in both left and right hands.

*Screenshot of training the model*

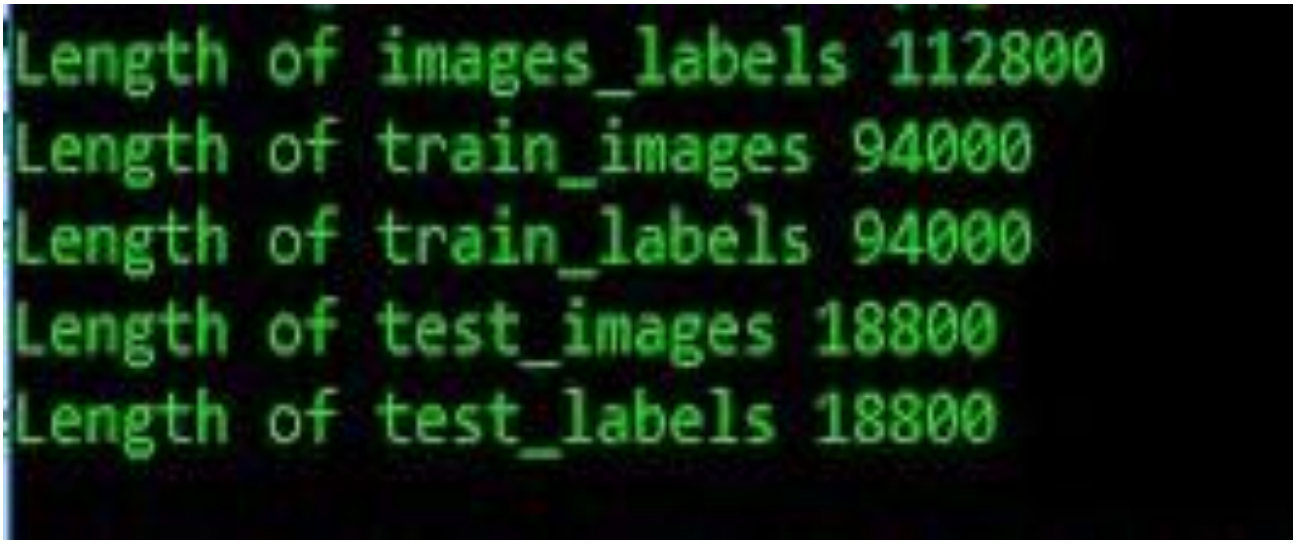


**Figure 8 Training and testing images**

Figure 8 represents the number of images in training set and the testing set. 5/6th of the entire dataset is used for training, remaining 1/6th of the dataset is used for testing. The total dataset consists of 44*2400 images of which 94000 images are used for training and remaining 18800 for testing.

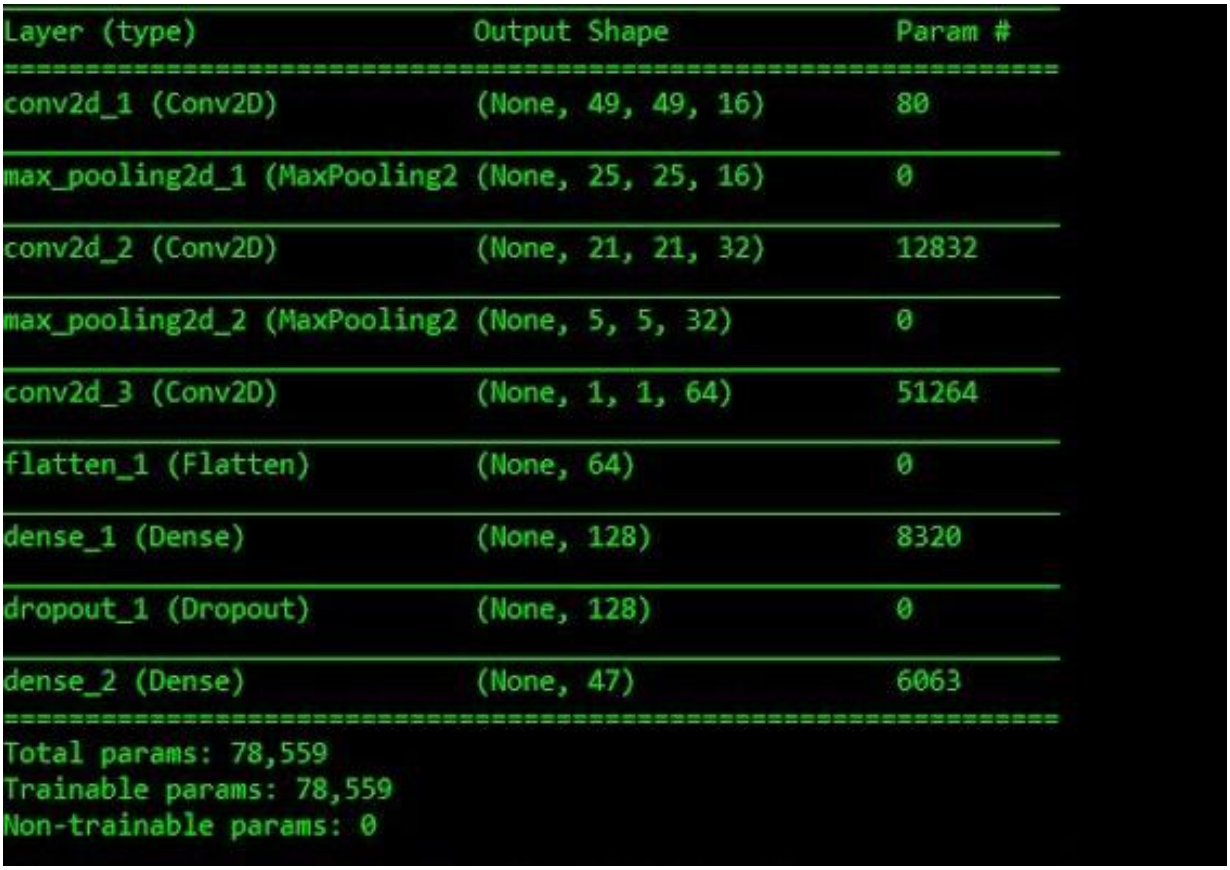


**Figure 9 Model Summary**

Figure 9 represents the model summary. We have 3 convolutional layers and 2 pooling layers which show the size reduction of the image with parameters being
captured. A flatten layer, with dropout of 20% of samples after which the final dropout layer gives the classification.

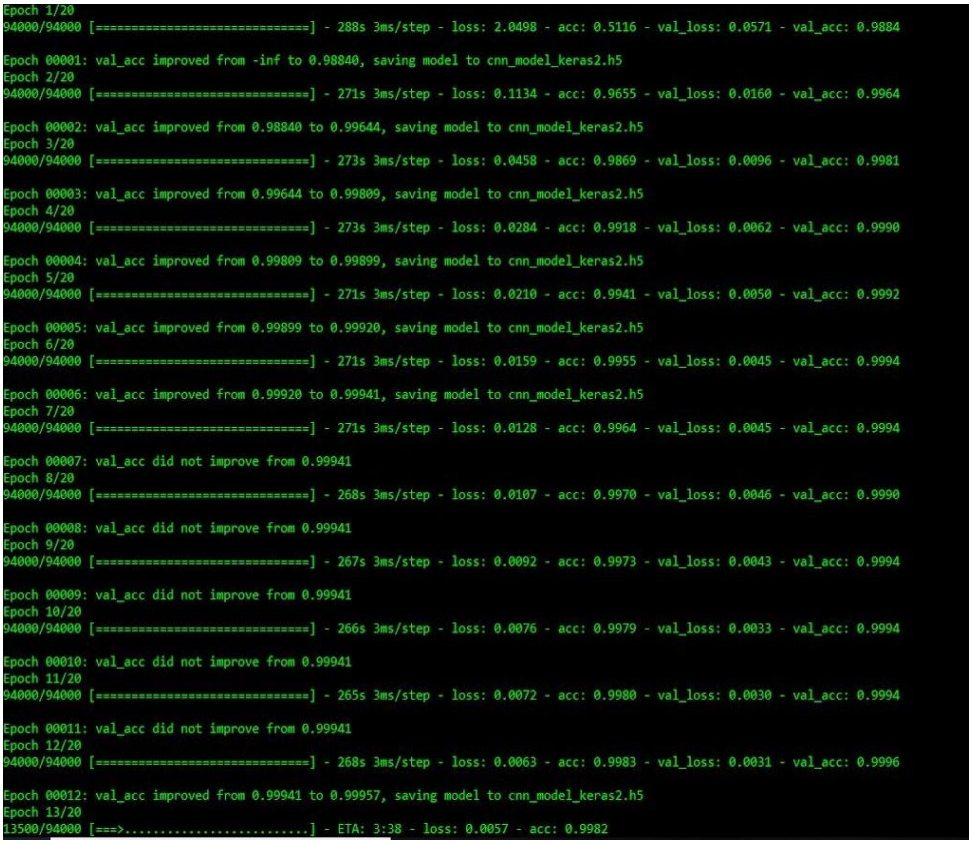


Figure 10 the training process

Figure 10 shows the training process of the CNN model. Each step gives the time taken to train, accuracy of testing and training. There are 20 epochs of training after which the model is saved to the file cnn_model_keras2.h5.

*Screenshot of dataset*

**Figure 11 Dataset**

Figure 11 This is the screenshot of 47 gestures created. A random image of each gesture is taken from the dataset for generating this image.

*Screenshot of recognizing gestures*

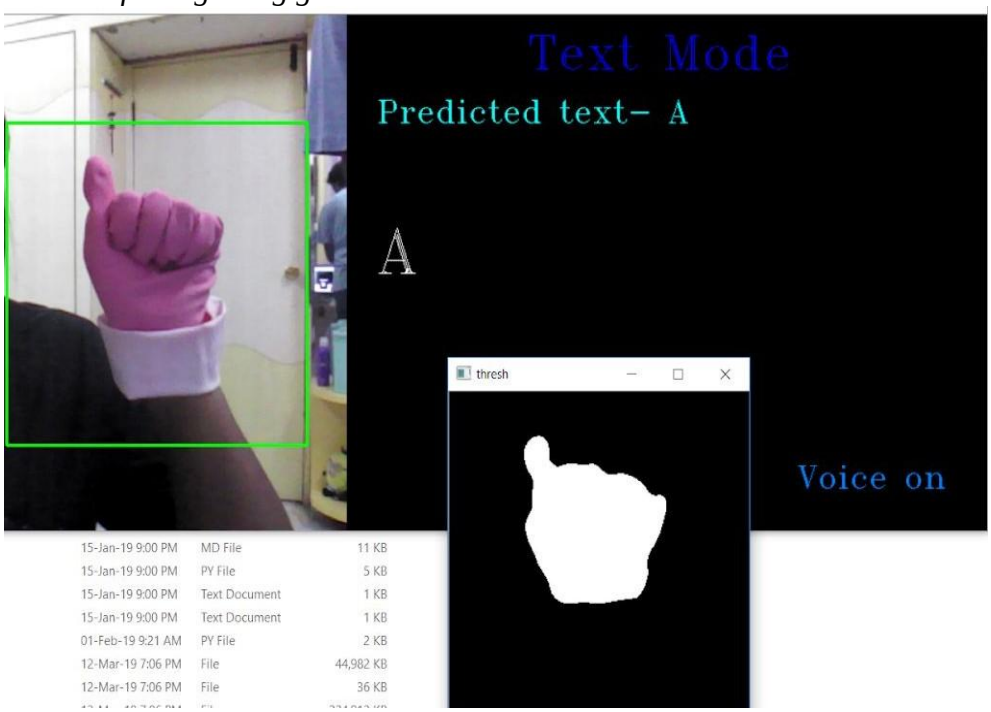


**Figure 12 Recognizing letter 'A'**

The Figure 12 shows the hand gesture which represents 'A' in ASL being predicted correctly. After the model captures 10 frames of the hand, the letter is predicted and it is translated to speech for the user.

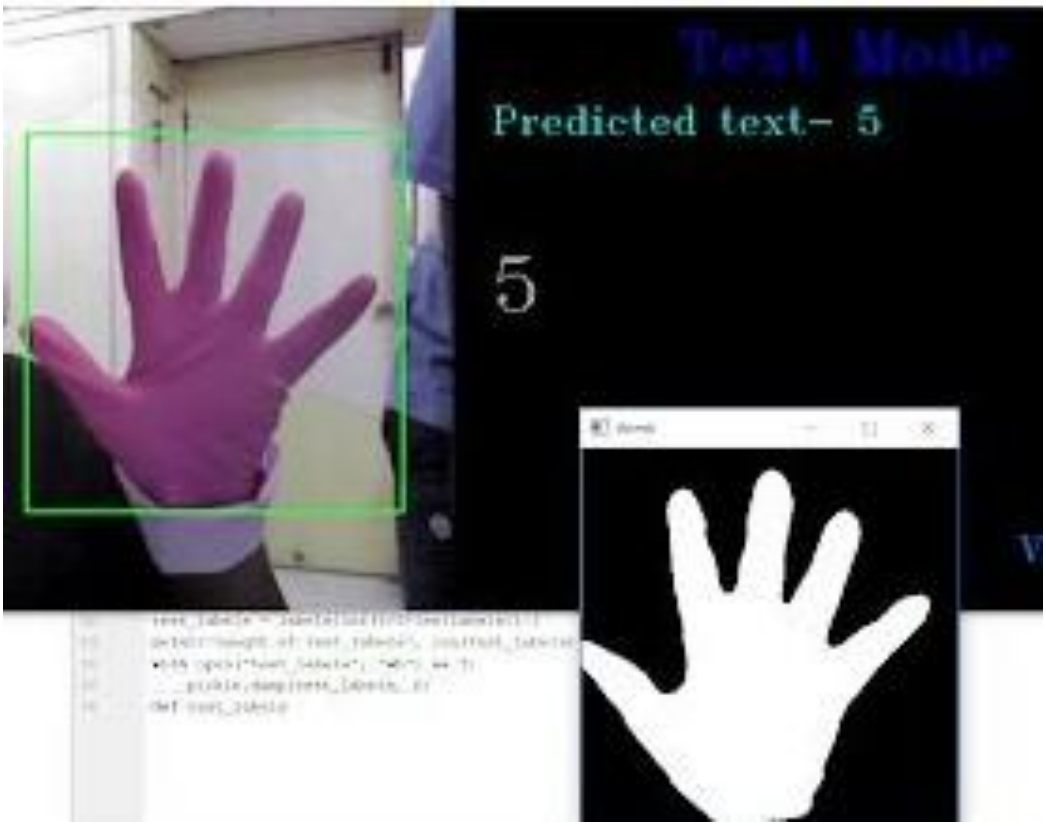


**Figure 13 recognizing number '5'**

The Figure 13 shows the hand gesture which represents number '5' being predicted correctly. After the model captures 10 frames of the hand, the number is predicted and it is translated to speech for the user.

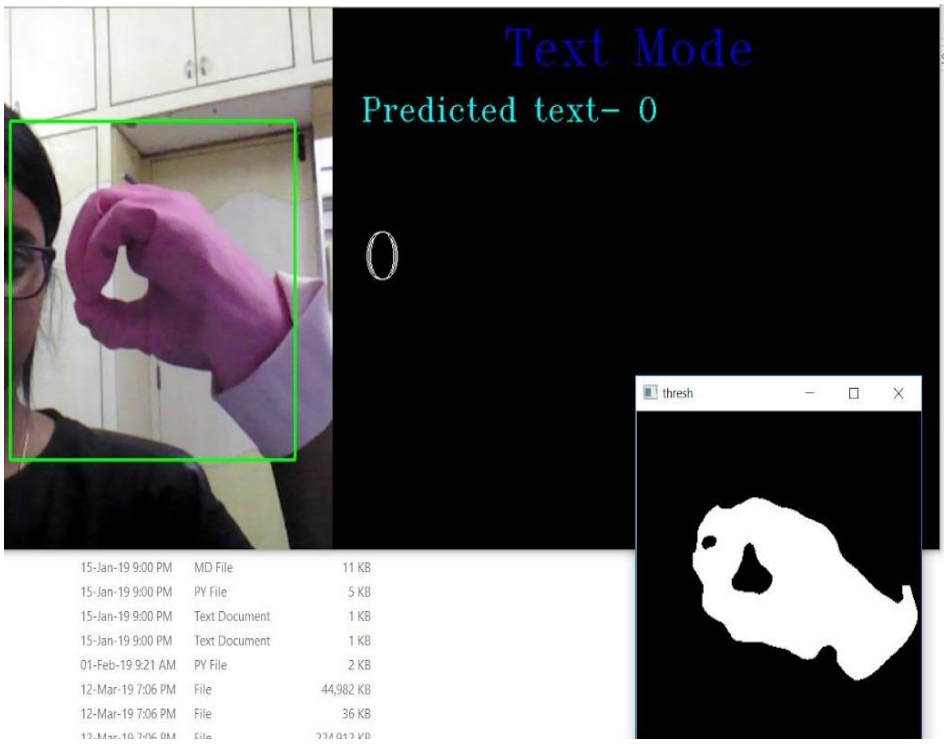


**Figure 14 Recognizing letter 'O'**

The Figure 14 shows the hand gesture which represents 'O' in ASL being predicted correctly. After the model captures 10 frames of the hand, the letter is predicted and it is translated to speech for the user.

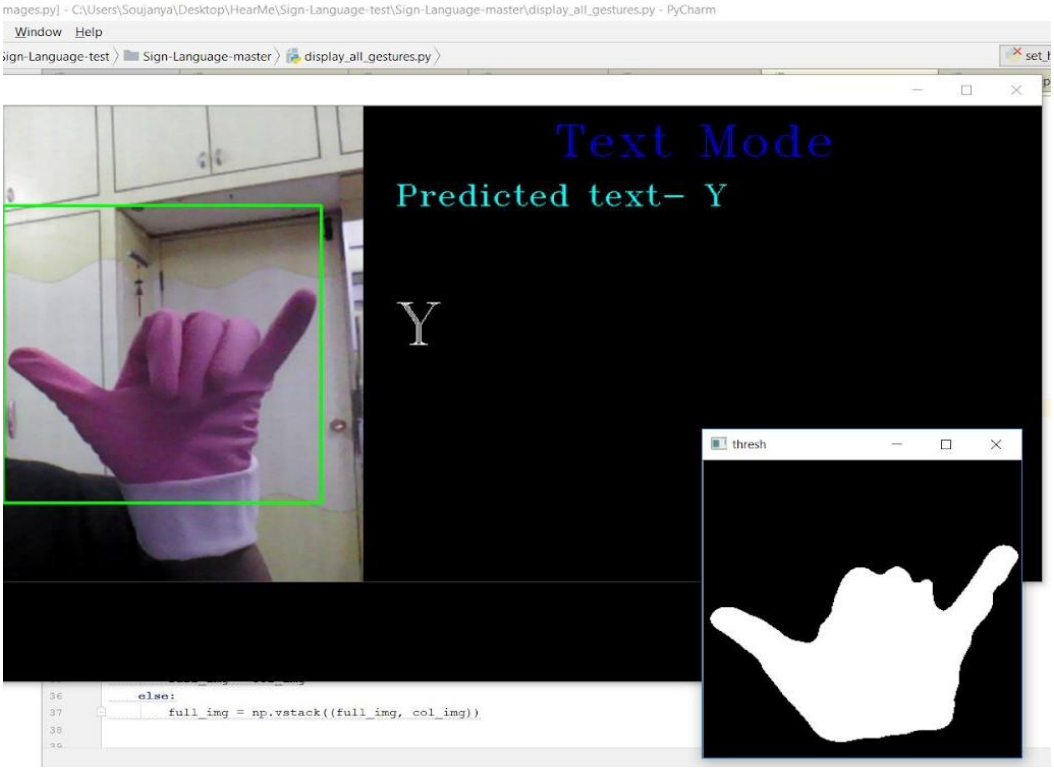


**Figure 15. Recognizing letter 'Y'**

The Figure 15 shows the hand gesture which represents 'Y' in ASL being predicted correctly. After the model captures 10 frames of the hand, the letter is predicted and it is translated to speech for the user.

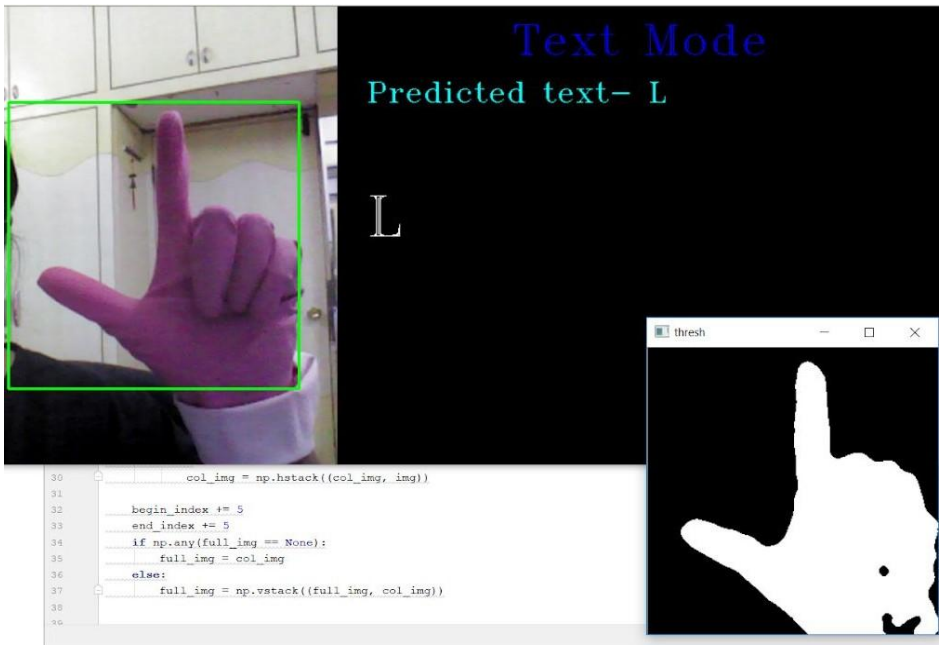


**Figure 16 Recognizing letter 'L'**

The Figure 16 shows the hand gesture which represents 'L' in ASL being predicted correctly. After the model captures 10 frames of the hand, the letter is predicted and it is translated to speech for the user. .

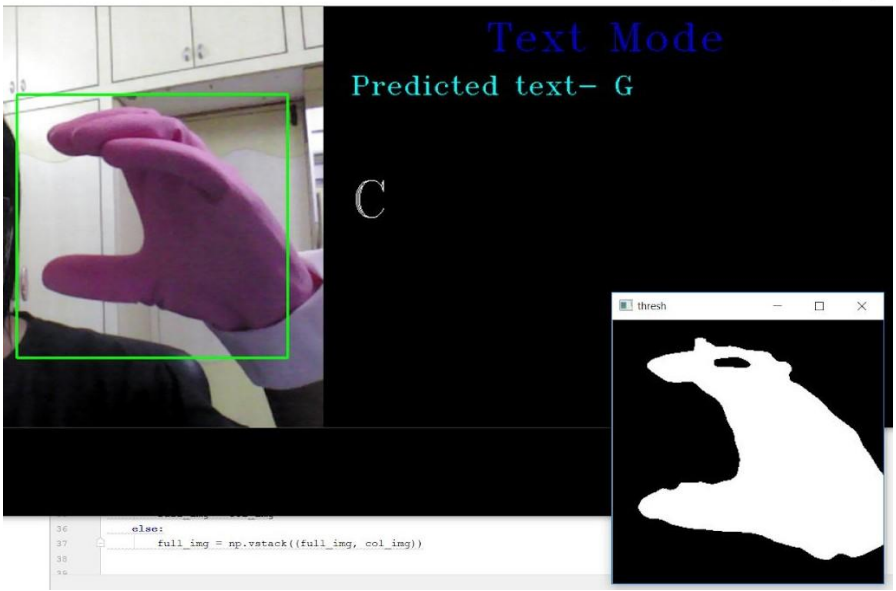


**Figure 16 Recognizing letter 'C'**

The Figure 16 shows the hand gesture which represents 'C' in ASL being predicted correctly. After the model captures 10 frames of the hand, the letter is predicted and it is translated to speech for the user.

## V. CONCLUSION AND FUTURE ENHANCEMENT

The CNN model fetched 99.4% accuracy while training and testing with the dataset. The problem of hand segmentation has been resolved with the usage of a colored hand glove and gave us better results. After working with different combinations of the number of convolution layers and pooling layers, 3 convolution layers along with 2 pooling layers fetched the highest accuracy. The disadvantage of training the new gestures is currently taking 80 minutes, hence the process of user adding a new gesture and training the model will be improved in the further stages of development. As the hand segmentation is dependent on the color of the hand, if the objects in the background match the skin color, it could distort the binarized threshold image. Due to similar gestures that exist in ASL, the final accuracy of classification depends on the environment and image processing techniques.

To overcome the above limitations, multiple cameras at different angles must be used to get the depth perception of the hand. The future work is to extend the desktop application to make it available in mobile devices for portability and ease of access. This project can also be extended to various sign language conventions.

## REFERENCES


1. Nikam, A. S., & Ambekar, A. G., "Sign language recognition using image based hand gesture recognition techniques."Online International Conference on Green Engineering and Technologies (IC-GET),2016.
2. Gurav, R. M., & Kadbe, P. K. "Real time finger tracking and contour detection for gesture recognition using OpenCV" International Conference on Industrial ,2015.
3. Panwar, M., "Hand Gesture based Interface for Aiding Visually Impaired", International Conference on Recent Advances in Computing and Software Systems, 2012.
4. G. R. S. Murthy & R. S. Jadon, "A REVIEW OF VISION BASED HAND GESTURES RECOGNITION."International Journal of Information Technology and Knowledge Management,2009.
5. Syed Atif Mehdi Yasir Niaz Khan, "SIGN LANGUAGE RECOGNITION USING SENSOR GLOVES." Proceedings of the 9th International Conference on Neural Information Processing (ICONIP),2002.
6. More, S. P., & Sattar, A. (2016). Hand gesture recognition system using image processing. 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT).
7. Kamal Preet Kour, Dr. Lini Mathew, "Sign Language Recognition Using Image Processing".International Journals of Advanced Research in Computer Science and Software Engineering.
8. Amit Kumar Gautam, Ajay Kaushik ,"American Sign Language Recognition System Using Image Processing Method ". International Journal on Computer Science and Engineering (IJCSE).
9. AlexKrizhevsky, IlyaSutskever, IlyaSutskever "ImageNet Classification with Deep Convolutional Neural Networks ".
10. Saad Albawi, Tareq Abed Mohammed "Understanding of Convolution Neural Networks". IEEE Conference-2016.


## AUTHORS PROFILE

Dr.E.Padmalatha completed research on classification of real time data streams .Working in Chaitanya Bharathi Institute of Technology it is a private engineering college as Assistant professor, published 30 research papers in different journals and 15 international conferences .Presently guiding the students in machine learning.

Dr. R. Ravinder Reddy have completed his B.Tech (CSE) and M.Tech (CSE) from JNTUH and completed his Ph.D (CSE) in 2016 from JNTUH. He has 15 years of teaching experience and published over 33 research papers in International Journals and Conferences of repute and has successfully completed many research projects. He has guided 11 PG scholars. As a facilitator for the learning process organized 3 STTPs/ Workshops /FDPs /SDPs, 13 keynote and invited talks. He is Principal Investigator for 3 Research and 2 Consultancy Project. He was the recipient of Best teacher award in CBIT 2018. Dr. Reddy is very active in Professional Societies and is a Life Member of ISTE & CSI.

Besides the Bachelor's degree in Computer Science, sound Hands-on experience on various blockchain platforms, Ethereum and Hyperledger Umbrella being the mains with the tag of certified blockchain supply chain expert.

**Divyarsha Koduri** did her Bachelor of Engineering in Computer Science from Chaitanya Bharathi Institiute of Technology, Hyderabad. She is pursuing her masters in Computer Science from University of Texas at Dallas with Data Science as her stream. Her passion lies in applying Data Science methods to real world data and integrating its results into applications.

**Anil Krishna Chintapalli** did his B.E. in Computer Science from Chaitanya Bharathi Institiute of Technology, Hyderabad. He is working as an Application Engineer primarily focusing in the field of Data Engineering. His passion lies in solving real world problems using programming.