International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 06 | June 2020 www.irjet.net p-ISSN: 2395-0072

Communication of Speech Impaired People with Home Automation

Divyanshu Mittal[1], Priyanshu Mehta[2]

[1]*Student, Dept. of Information Technology, Bhagwan Parshuram Institute of Technology, Delhi, India*
[2]*Student, Dept. of Information Technology, Bhagwan Parshuram Institute of Technology, Delhi, India*

Abstract - *All over the world, various challenges are experienced by speech and hearing-impaired people in expressing themselves to other people. Gestures play a major role in the daily activities of a human life. Between all the gestures performed, hand gestures play an important role which express more in less time. The aim is to develop a software that can mediate the process of proficiently recognizing an image, of the gesture performed in front of the camera, with predictably high speed and accuracy. The system is interactive enough to provide aid to speech impaired people, even with no prior knowledge of sign language, able to use home automation technology without using any vocals. For this purpose, 28 different gestures have been considered where 2400 images per gesture are present. The approach of this model involves implementation of Deep Learning and Convolutional Neural Network as it is an excellent framework for robust gesture recognition. The idea of the convolution layer is to transform the input image in order to extract features to distinguish them correctly. Upon implementation of CNN, the model is able to identify the gesture, displays its corresponding text and convert it into voice message to be used as command for performing desired action.*

Key Words: **Gesture Recognition, Convolutional Neural Networks, Text to Speech, Sign Language Prediction, Google Assistant, TensorFlow, Rectified Linear Unit**

1. INTRODUCTION

Speech impaired people over the world use sign language for communication with one another. Sign language is a viewable language which consists of manual actions that are either facial or require body movements. Sign language is not a universal language, and different sign languages are used in different countries. Even there are some countries like the UK, the USA, Belgium or India that may have more than one sign language. Hundreds of sign languages are in used around the world, for instance, Japanese Sign Language, British Sign Language (BSL), Spanish Sign Language, Turkish Sign Language.

In spite of the fact that technology has been a boon to the world, a section of the society is unable to get a handful of it. It is important that certain systems are designed in order to support specially-abled people who lack behind otherwise. The speech or hearing-impaired people cannot communicate the way a normal person can and thus implement the 'sign language' through the means of actions, hand shapes, hand movements and location of hands with respect to the body.

The system should be able to proficiently recognize the image of gesture upon displaying in front of a camera, interpret and translate it into readable text. Upon translation, the aim is to convert it into a voice signal that can be utilized to operate home automation. This can allow devices such as computers, security, and camera systems to be controlled remotely from any room as well as from other locations by a smartphone or internet. With home automation, manual interference is minimized as most of the appliances are controlled smartly by a computer.

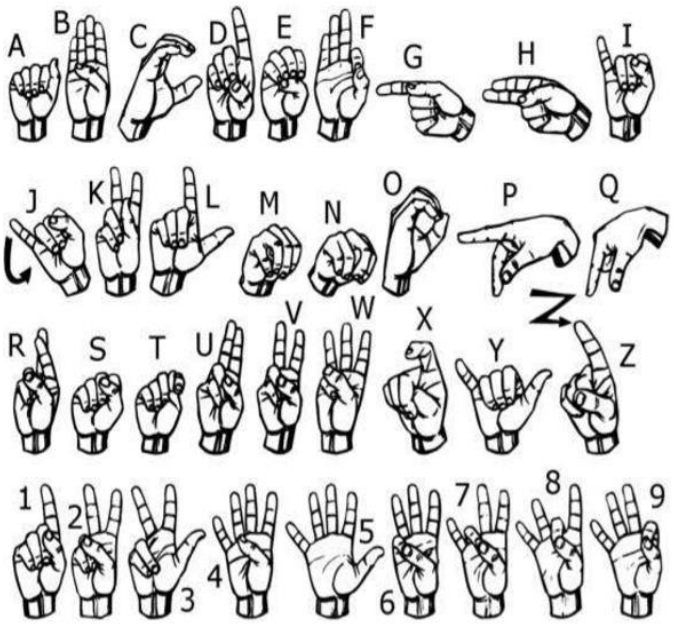

Fig -1: Finger spelling American Sign Language [1]

The alphabet level sign language is easy to understand but at the same time it is cumbersome to use in day to day life

of an impaired person, this is why word level sign language was developed as it lets the gates to open and effective communication between the participants or in other words it helps in effective transmission of message from the sender to receiver. Sign language recognition is a problem that needs to be tackled differently in each region, since new movements or hand shapes or combinations thereof require new training data, and possibly involve new challenges which needs to be overcome before stepping into gesture recognition.

Gesture recognition [2] can be seen as a way for disabled to interact with others with the aid of computers. Also, it builds a richer environment between machines and humans which is much better than primitive text user interface, which still limits inputs through hardware components manually. Using this concept, it is now possible to point a finger at a point and it will move accordingly. Gesture recognition has been accomplished by experimenters using different techniques. One that is implemented in this research is Convolutional Neural Networks (CNN). It is one of the most flexible, manipulative and exemplary techniques available. A CNN [3] is made up of neurons that have learnable weights and biases. Each and every neuron in the network receives some inputs, based upon the system's framework, then performs a dot product and optionally follows it with a non-linearity. These are special kind of multi-layer neural networks, designed to recognize visual patterns directly from pixel images with minimal pre-processing.

2. LITERATURE REVIEW

After observing the unique approaches, multiple researchers have acquired the undermentioned insights on 'Hand gesture to Speech conversion' and 'Home automation'. Some investigated various techniques to follow on a pre-determined sign language dataset whereas some generated their own set of signs and symbols.

Dhruv Vaghela, Ankit Vaity, Rajvi Makwana, Afreen Shaikh, Chanda Chouhan [4] proposed a system based on live feeding sign gestures through a webcam which deploys TensorFlow and Keras package of Python3 with Heroku (cloud platform). Aiming to get the desired results, after training and filtering the data collected, CNN model is used to detect the gesture which is afterwards shown on screen and spoken through a speaker.

The work of Prof. R. R. Itkarkar and Dr. Anil V Nandi [5] involves conversion of hand gesture into speech by using MATLAB. The camera feeds the images as input which are later converted to binary image. As the next step Matlab algorithm is used to detect the gesture from the binary and the pre-recorded gesture's name and speech is played.

Rutuja et al. [6] researched and evaluated the Hand Gesture Recognition System Using CNN by exhibiting a comparative analysis of three approaches: CNN, KNN, ANN. The prepossessing outcome of the researchers manifest that CNN is the most suitable approach for recognition purposes. The overall accuracy of the system developed by them was 94.32%.

In the research of Aki Kunikoshi, Yu Qiao, Nobuaki Minematsu, Keikichi Hirose [7], 'Speech generation from hand gestures based on space mapping' was performed. A glove which senses the spatial actions of the hand and simultaneously speaks out the meaning of whatever gesture is performed. This framework optimized the time taken in the translation of hand gesture into speech and allowed the user to prefer this over synthesizer used for communication between abled and disabled people.

The model of Priyanka D. Hatwar, Neha A. Wahile, Isha M. Padiya [8] incorporated home automation system based on gesture recognition system with software as well as hardware components. Aiming towards developing a helping hand for the ageing population the model senses the gesture and thus, predict advanced speech output which subsequently turn on/off various appliances in the home.

The proposed research of Ali Hussein et al. [9] depicted a Smart Home Design for Disabled People based on Neural Networks. The system necessitates capabilities which can aid people according to their disabilities. It encourages impaired people to operate smart homes effortlessly without troubling them to move from their place.

3. METHODOLOGY

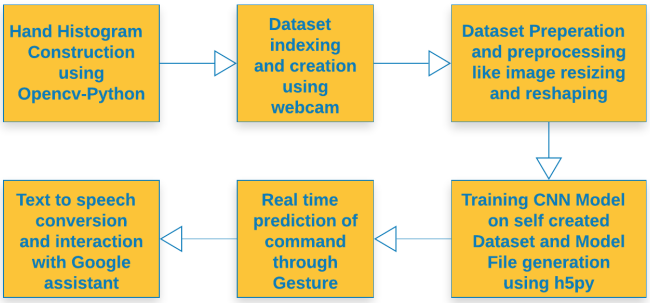

Fig -2: Process flow diagram

3.1 Data Acquisition

The dataset is a self-created hand gesture-based image data which has been curated with close consideration of the Indian Sign Language. The database is accumulation of 67,200 images. It comprises of 2400 images of each of the 28 signs that have been considered effective for the purpose of operating home automation. User can add personalized gestures according to one's use case and operations of home automation.

For creating dataset, it is important to set a standard for recording and saving new gestures, therefore we've defined a 300x300 window that records hand's histogram in a particular lighting condition. Gestures can be generated and stored one's a good histogram has been generated.

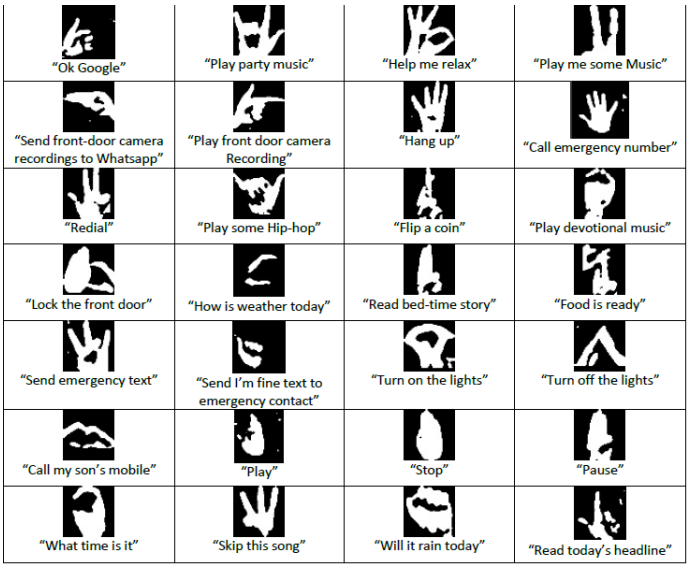

Fig -3: Gestures used in demonstration

3.2 Pre-Processing

Since there are multiple copies of each image, we introduce image rotation and flipping of original gesture to make database variegated as well as introduce different images to avoid over fitting of data.

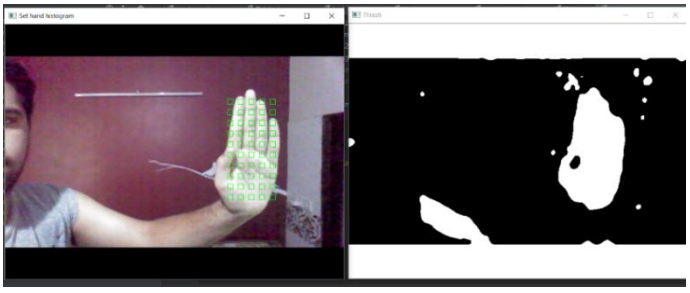

Fig -4: Histogram creation

3.3 Train Test Split

We have used the golden ratio of 80-20 to divide our dataset in training and test data. Out of 2400 images for each gesture, we've used 1920 images as training data and the test-train data is randomly split to decrease generalization.

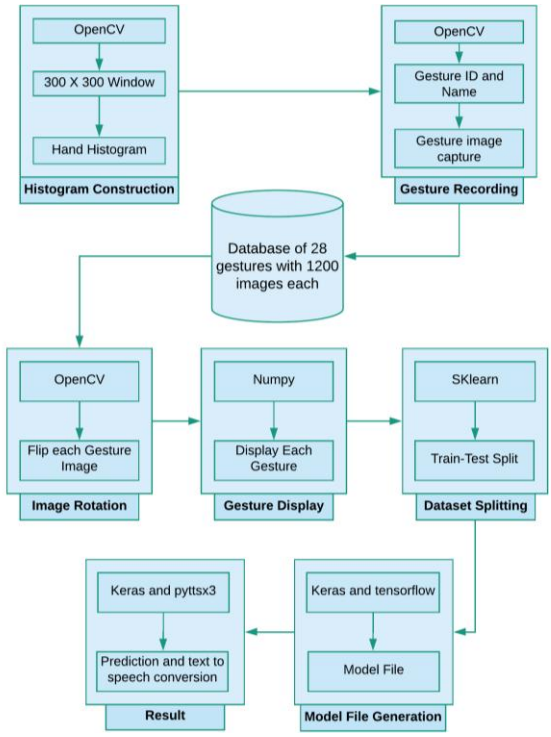

Fig -5: Data flow model

3.4 Model Generation

Convolutional Neural Network (CNN) is a well-known neural network for treating images and positional features. CNN has an outstanding ability of identifying relations and recognizing patterns with ease and can cope with any translational or rotational changes that may occur in the images. We have utilized CNN layers for training data by using ReLu activation function and an amalgam of different layers of CNN from Keras. We've then generated a model file used in real-time gesture prediction.

Fig -6: Predicted Gesture "Play me some music"

3.5 Communication with Home Automation

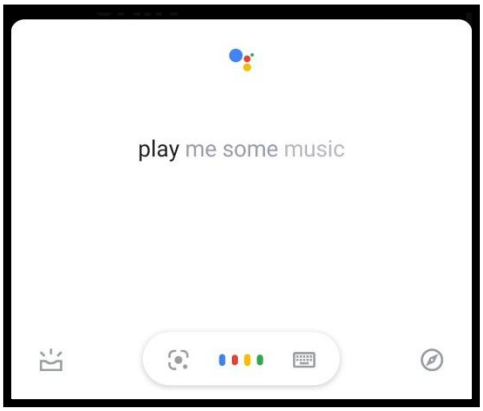

Fig -7: Audio captured by Google Assistant

After successful prediction of the hand gesture, it's communication with any home automation system or virtual assistant system depends upon voice output and not textual output. Therefore, text to speech conversion for English language has been utilized to convert text output generated by the model to speech output which acts as a command for home automation system.

4. RESULT

In order to recognize hand gesture and predict its meaning, the model needs to understand dimensional features and then generate required text output. For the purpose of instructing home automation systems, we need to convert this text-based output into an audio stream which can then act as an input to virtual assistants and can get work done with ease.

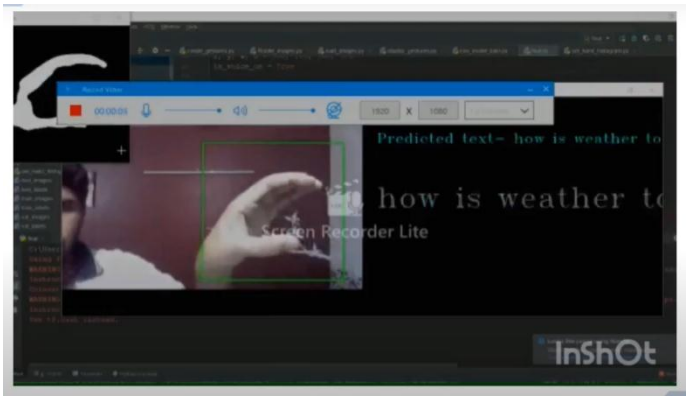

Fig -8: Predicted Gesture "How is weather today"

During model generation we recorded training accuracy of 93.05%, because of which there are moments when action is not correctly recognized in real-time scenario. Hand gesture recognition is a vast field with extensive research and our idea is a small contribution.

Fig -9: Result generated for "How is weather today" by Google Assistant

5. CONCLUSIONS

Hand gestures are a powerful way for human communication, with lots of potential applications in human computer interaction. Vision based hand gesture recognition techniques have many proven advantages compared with traditional devices.

We've presented a vision-based system which is able to interpret isolated hand gestures from self-created dataset and translate it to text and further speech to interact with an Automation Systems.

Conventional programming makes it difficult to build solutions for computer vision because of generalization as an outcome of which they are unable to identify variations in images. Image classification by implementing CNN with TensorFlow yielded appreciable results. We obtained training accuracy of 93.05%. In future, extended work may include support to different lighting conditions and incorporate feedback mechanism to make model more robust. Further work may also encompass real-time recognition of moving gestures, which will help in making the system verbose.

REFERENCES

[1] National Institute on Deafness and Other Communication Disorders (NIDCD) https://www.nidcd.nih.gov/health/american-sign-language Retrieved- Nov 2019.

[2] Sushmita Mitra and Tinku Acharya, "Gesture Recognition: A Survey", IEEE Transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 37, no. 3, May 2007, pp. 311-324

[3] Jiang, D., Li, G., Sun, Y., Kong, J. and Tao, B., "Gesture recognition based on skeletonization algorithm and CNN with ASL database", Multimedia Tools and Applications (Springer), 2018

[4] Dhruv Vaghela, Ankit Vaity, Rajvi Makwana, Afreen Shaikh and Chanda Chouhan, "Hand Gesture Recognition to Speech Conversion", IOSR Journal of Engineering (IOSR JEN) vol. 7, INFT, 2019, pp. 21-25.

[5] Prof. R. R. Itkarkar and Dr. Anil V Nandi, "Hand Gesture to Speech Conversion using Matlab", 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), IEEE

[6] Rutuja J., Aishwarya J., Samarth S., Sulaxmi R. and Mrunalinee P, "Hand Gesture Recognition System Using Convolutional Neural Networks", International Research Journal of Engineering and Technology, vol. 06, no. 04, 2019, pp. 4508-4514.

[7] Aki Kunikoshi, Yu Qiao, Nobuaki Minematsu and Keikichi Hirose, "Speech Generation from Hand Gestures Based on Space Mapping", 10th Annual Conference of the International Speech Communication Association (ISCA), Brighton, UK, 2009, pp. 308-311.

[8] Priyanka D. Hatwar, Neha A. Wahile and Isha M. Padiya, "Home Automation System Based on Gesture Recognition System", International Journal of Emerging Technologies in Engineering Research (IJETER) Volume 5, Issue 3, 2017, pp. 48-53.

[9] Ali HUSSEIN, Mehdi ADDA, Mirna ATIEH and Walid FAHS, "Smart Home Design for Disabled People based on Neural Networks", The 5th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN), 2014, pp. 117-126