International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 03 | Mar 2020 www.irjet.net p-ISSN: 2395-0072

Hand Gestures Recognition using Deep Learning

Dharmik Bhanushali[1], Dipanshu Someshwar[2], Vismay Chaudhari[3], Ms. Bhargavi Dalal[4]

[1,2,3]B.E. Student, Shah & Anchor Kutchhi Engineering College, Mumbai, Maharashtra, India
[4]Assistant Professor, Dept. of IT Engineering Shah & Anchor Kutchhi Engineering College, Mumbai, Maharashtra, India

---

Abstract - *Digital assistant devices such as Alexa, Google Home, Siri and Cortana are increasingly prevalent. These devices listen for every command, their platforms know where you are, and they have a deep and evolving understanding of your preferences. Unlike normal people, dumb and deaf people find it hard to use a personal assistant as it is audio based. The main objective is to build a system that allows deaf-mutes to use personal assistants easily without facing any difficulty. The idea is to create an interface that allows voice controlled devices to accept inputs also in form of hand gestures and generate output in text format by making use of Deep Learning Modules.*

Key Words: Deaf-mutes; deep learning; digital assistants; hand gestures; Machine learning.

1. INTRODUCTION

People with hearing or speaking problems have personal and social limitations. Such limitations will greatly affect their ability to perform everyday tasks like shopping and travel independently, using modern technology- tools. The identification of positive and successful ways to enable people with a hearing disability to perform group activities may help them overcome their disabilities and retain or enhance their independence. This project forms part of ongoing work aimed at using gesture recognition technology in the design of a program that can help empower people with hearing or speaking disabilities to enhance social and communication skills and the quality of life with other people. Use of such technology would not only boost their individual life experience but also help deaf-mutes to get more opportunities and considerations in their professional life.

2. METHODOLOGY

Machine learning is a scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions, relying on patterns and inference instead. Algorithms like Convolutional Neural Network and K-Nearest Neighbours can be used according to the requirement of the project. In the earlier versions, the hand gestures performed and recognized were static which might be a bit outdated and the training images were clicked by the user itself which can be quite a task for the user and may have some difficulties as the process becomes user-oriented. This idea can be further improved by bringing in real time gesture recognition with a pre-trained dataset which can provide more efficiency and make user's experience much better compared to static gesture recognition.

The use of Deep Learning over Machine Learning is done because Machine Learning shows output accuracy in form of 0 or 1, whereas Deep Learning provides better numerical accuracy results that ranges from 0 to 1 [6]. Since the proposed project demands high need of accuracy in Classification Algorithms it is better to make use of Deep Learning. Just like Data Classification is the most important aspect of Data Science, similarly Image Classification plays a very big role in Computer Vision. Image Classification first deals with image pre-processing, then it executes image segmentation and then it deals with key feature extraction that helps for finding uniqueness in different trained images and lastly executes matching identification [5].

TensorFlow computations are purely based on use of Tensors where tensors can be considered as vector or matrix which is used to represent all kinds of data. All the operations in TensorFlow takes place inside a Graph where each operations are termed as op node and are interconnected. We can manually add data to TensorFlow but in order to make use of TensorFlow with Deep Learning we make use of Pipeline to add data to TensorFlow for Training [4].

Basically Image Processing will help Computer Vision for extracting useful information out of trained images. Then Artifical Nueral Network Algorithm which comes under Machine Learning will be used for Classification purpose to classify the current hand gesture with series of trained images and accordingly produce output on basis of Accuracy.

2.1. Stepwise procedure:

1) Series of hand gestures will be performed in front of the webcam.
2) Each sign then will be converted to text format.
3) The text will be converted to audio using a python module.
4) The audio will act as an input to the assistant.

5) The query will be processed by the assistant.
6) The response by the assistant will be in audio format.
7) This response will be converted into text using a python module which will be helpful for people with hearing disabilities.
8) The text will be displayed accordingly as the response to the input.

2.2 .Block diagram:

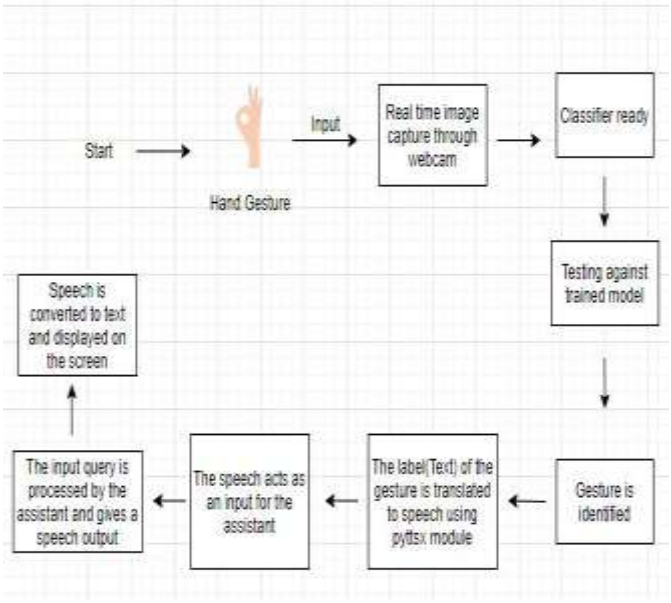

Fig 1: Block Diagram

3. LITERATURE REVIEW

Till date, majority of digital assistants are voice automated thus restricting the use to specific customer who may have hearing disabilities. The current modules which are developed with similar intentions of this project lacks in real-time conversion of hand gestures as they first capture the image then processes it which increases the overall time to get the output [1].This leads to need of real time gesture recognition system. Few similar modules fail to process series of words and are limited to fewer number of words while converting to audio format [4]. This leads to need of better design of text to speech systems or use of better modules for the same. No matter how good is the performance of Machine Learning modules but proper user interface is equally important which was not found in many such modules. This leads to need of designing a more user friendly system. Initially technology was developed for Sign language recognition based on Hardware (for instance say Cyber glove) [3]. This leads to use of wearing hardware equipment all the time, which was again not practically feasible.

4. PROPOSED WORK

This project specifically concentrates on utilizing digital assistants (say Alexa) which is an audio based Digital Assistant in order to replace speech recognition with gesture recognition technique in orderly manner. Our proposed system requires a neural network to interpret gestures (i.e. Conversion of video to text), text to speech system to speak the interpreted sign to digital assistant, speech to text system to transcribe the response from assistant to the user, a laptop, interface that ties all this together for a flawless experience. It will first undergo Data division using datasets and Augmentation techniques to increase the size of dataset [2]. Secondly Classification will be done using deep convolutional network [2].

After successful completion of the proposed module, it will be useful for a wider range of users including deaf-mutes. The underlying design makes it capable for processing real-time gestures and comparatively at a faster pace. The current module intends to convert comparatively more number of words to audio format which then becomes actual input to Digital Assistants. This design focuses equally on the Interfacing part for overall better experience of user. The proposed system is Vision based approach in which user does not need to wear any sort of hardware components.

4.1 Advantages of the Approach

- The system will help in achieving efficient results in minimum time.
- Real time image capture and gesture recognition.
- More accurate performance than previous systems.
- Efficient and better use of Text-to-Speech and Speech-to-Text modules, thus giving smooth experience.
- No need of hardware equipment attached to user's body parts at any point of time.
- A user friendly mode of display for easier understanding of the user and for flexible usage.

5. CONCLUSION

This project reviews the current state-of-use of Digital assistants. A different method is proposed to access or to communicate with audio based digital assistants. In this project various regression and classification models will be used for increasing the classification accuracy. Overall an interface will be created which will help deaf-mutes to use the voice automated devices using hand gestures or sign language for the input and also get outputs in form of text displayed on screen for those with hearing disabilities.

6. REFERENCES

[1] Yusnita, L., Rosalina, R., Roestam, R. and Wahyu, R., 2017. Implementation of Real-Time Static Hand Gesture Recognition Using Artificial Neural Network. CommIT (Communication and Information Technology) Journal, 11(2), p.85.

[2] Rathi, P., Kuwar Gupta, R., Agarwal, S. and Shukla, A., 2020. Sign Language Recognition Using ResNet50 Deep Neural Network Architecture. SSRN Electronic Journal

[3] V. Adithya, P. R. Vinod and U. Gopalakrishnan, "Artificial neural network based method for Indian sign language recognition," *2013 IEEE Conference on Information & Communication Technologies*, Thuckalay, Tamil Nadu, India, 2013, pp. 1080-1085.

[4] Guru99.com. 2020. Tensorflow Image Classification: CNN(Convolutional Neural Network). [online] Available at: <https://www.guru99.com/convnet-tensorflow-image-classification.html>.

[5] Guo, T., Dong, J., Li, H. and Gao, Y., 2017. Simple Convolutional Neural Network on Image Classification. IEEE 2nd International Conference on Big Data Analytics, pp.1-2.

[6] Ertham, F. and Aydin, G., 2017. Data Classification with Deep Learning using Tensorflow. IEEE 2nd International Conference on Computer Science and Engineering, pp.757-757.