iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** I **Month of publication:** January 2021

DOI: https://doi.org/10.22214/ijraset.2021.32813

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue I Jan 2021- Available at www.ijraset.com

Implementation of Virtual Assistant with Sign Language using Deep Learning and TensorFlow

Vani Valsaraj
Department of Computer Science and Engineering, APJ Abdul Kalam Technological University

Abstract: ***Sign languages are languages that are use usual manual modality to convey meaning. Gestures are the form of non-verbal communication in which visible body action are used to communicate. This project is all about the system interface developed that allows deaf and mutes to make use of various voices with the help of sign language. Majority of virtual assistants work on the basis of audio input and gives audio output only but my project we will build a speak engine and on successful identification of hand gesture we will trigger the speak engine. It make use of the concept of deep learning, CNN tensor flow library etc. Finally the result will be in the form of text and we will use that text to speak out.***
Keywords: Sign language, virtual assistant, TensorFlow, Deep learning, Python

I. INTRODUCTION

Gesture means it is form of body motions mainly hand, leg etc. Deaf and dumb community perform sign language. This community use sign language for their communication when broadcasting audio is impossible or typing and writing is difficult but there is the vision possibility. At that sign language is the only way of exchanging information between people. Normally sign language is used by everyone when they do not want to speak but this is only way of communication for deaf and dumb community. Sign languages have the same meaning that of spoken language.

Sign languages are of two types they are isolated and continuous sign language. Isolated sign language means single gesture mainly of single words whereas continuous sign language means it is a sequence of gesture having meaningful sentences. In this project we will be performing isolated American Sign Language recognition techniques.

II. LITERATURE SURVEY

In [1] a real time dynamic hand gesture system performed where eleven hand gesture recognized. Author use YCbCr color space for detecting skin color.This system is mainly used in the application of Hcl. In [2] Author proposes a smart gloves which convert the sign language to speech output. Flex sensors are placed between hands In [3] Author uses softMax classifier for the neural network purpose and finally results show most accurate classification.[4] Author build simple cnn on image and mainly benchmarking dataset are used and they verify the shallow network also have good recognition.[5] goal is Artificial intelligence(AI) which realize natural dialogue between humans and machines. Here author uses multi-model systems which have more combined input models.

III.PROPOSED SYSTEM

In the proposed system for gesture recognition we are using image processing and computer vision .Gesture recognition enables computer to understand human action and also act as an interpreter between computer and human. This could provide potential to human to interact naturally with computers without any physical contact of mechanical devices. Concepts of CNN and tensorflow libraries are used.CNN is nothing but deep learning algorithm that is capable of assigning biases and weights to different objects in an images and on basis of same it can differentiate one images from another.

It consist of different processing layers of image classification and it is designed with the mean of representing functioning of neurons in human brains.

The best part of using tensorflow library is that it is an open source with lot of pre designed model useful in deep learning. Tensorflow in our system helps in training the model using provided dataset. It is help in classifying and identifying real time hand gesture help in processing more information a spot more patterns.

The most basic explanation of workflow of the system goes as follows - A hand gesture is performed in front of the webcam. This sign gesture is converted to text and the text output is converted to audio and is served as an input to the assistant. The assistant processes the question and responds in audio format. This audio format is converted to text output. The text output will be then displayed on the display screen.

©IJRASET: All Rights are Reserved

A. Raw Data

The very first step is providing the raw material or dataset. Once the extraction of frame done from the raw dataset the next phase is preprocessing. This is done by key frame extraction using gradient method and after preprocessing next comes the feature extraction is done.

B. Data Preprocessing

Data preprocessing is used to convert the raw data into clear dataset. Key frame extraction is powerful tool that is implements video content by detecting a set of summary key frames to represent video sequences.

C. Feature extraction

Features are extracted from the dataset using key gradient based key frame extraction technique these key frames are useful for splitting continuous languages gestures into sequences for sign as for removing uninformative frames. After splitting of gestures features of preprocessed gestures are extracted using orientation histogram is done.

D. Training

The last one is training that is recognition is done using distance techniques finally output will be formed. Objective score is that summarizes the relative difference between two objects in a problem domain. Basically the model will be trained more number of times for the same set of labels, higher is the success rate. But it is required to keep a note that any changes made with the labels folder before training will lead to the system that is being trained for the very first time. In simple terms if it is required to make any changes in the labels folder that is adding new labels or replacing the existing labels, the model will need to be trained again from the beginning. It was observed that for training of around 15 labels on an average configured system, it takes about 12-15 hours straight of model training for the first time. However retraining of same set of labels requires comparatively lesser amount of time.

E. Output Generation

After training finally the output will be formed in the form of text and we will use that text to speak out. The entire process can be repeated multiple number of times. The hand gesture will be only recognized if they meet the trained dataset standard.

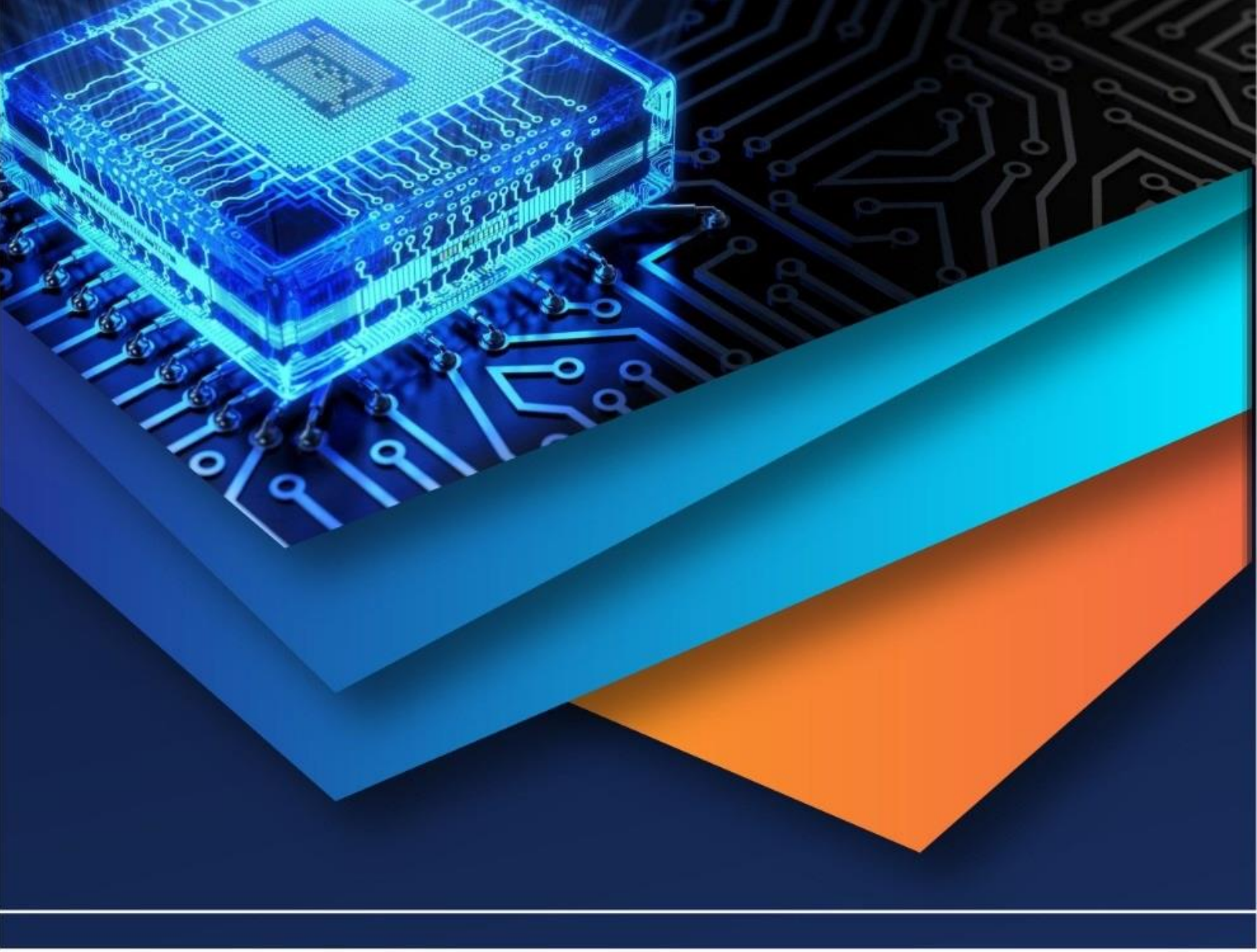
IV. CONCLUSIONS

The designed system was successfully able to capture Hand Gestures using the integrated Web Camera and process and convert into text format and display it onto the Input frame and then converted into Audio format. The audio output was being successfully converted into Text format and displayed on the screen. Basically the system can be considered as a boon to people with hearing disabilities or speaking disabilities or both at the same time. These systems would not only bring technology into their Personal lives but also give rise of opportunities in their professional life.

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. Aditya, P. R. Vinod and U. Gopalakrishnan, "Artificial neural network based method for Indian sign language recognition," 2013 IEEE Conference on Information & Communication Technologies, 2013
[4] Tensorflow Image Classification: CNN(Convolutional Neural Network) Guo, T., Dong, J., Li, H. and Gao, Y.,
[5] Lai, H. and Lai, H., 2014. Real-Time Dynamic Hand Gesture Recognition. International Symposium on Computer, Consumer and Control, pp.658-66

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)