International Journal of Research in Advent Technology (IJRAT) Special Issue
E-ISSN: 2321-9637
Available online at www.ijrat.org
National Conference on "Role of Information Technology in Social Innovations"
26th & 27th February 2019

An Android Application For Interpreting Sign Language Into Text Format Using Deep Learning

Mr. Niraj A. Haval[1] , Kedar V. Kulkarni[2] ,Omkar G. Patil[3] ,Aniket M. Mangalekar[4] ,
Paul P. Panhalkar[5] ,Prithviraj P. Badiger[6]
Assistant Professor[1]
Computer Science and Engineering D. Y. Patil College of Engineering & Technology Kolhapur, India[1,2,3,4,5,6.]
prithvirajbadiger11@gmail.com ,kedar.kulkarni626@gmail.com ,patilomkar6024@gmai.com, mangalekaraniket@gmail.com, panhalkarpaul@gmail.com, nirajhaval32@gmail.com

Abstract—Communication is the main channel between people to communicate with each other. In the recent years, there has been rapid increase in the number of deaf and dumb victims due to birth defects, accidents and oral diseases. Since deaf and dumb people cannot communicate with normal person so they have to depend on some sort of visual communication. There are many languages spoken all around the world and interpreted. "Special people", that is people who have difficulty in speaking and hearing "The dumb" and "The deaf" people respectively find it difficult to understand what exactly the other person is trying to express and so with the deaf people. Now a day's smartphone is affordable to every middle class person in India. Sometimes people interpret these messages wrongly either through sign language or through lip reading or lip sync. Our system is made in such a way to help these specially challenged people to hold equal part in the society.

Keywords—Sign Language, deep learning, convolutional neural net, Google, MobileNet SSD, TensorFlow.

1. INTRODUCTION

Nowadays we always hear about new technology that improves our lifestyle, which makes our life easier. Technology has revolutionized the human mankind. Human race has put a gear in technology and they are not in a mood to move the pedals away from this gear. There is huge research on various technology sector such as Artificial Intelligence, Smart phones and many more. This research lead to new inventions and making one's life easier. But there has been a very less research for Deaf and Dumb people. This topic has get less attention as compared to other sectors. The Main challenges that this special person facing is the communication gap between special person and normal person. Deaf and Dumb people always find difficulties to communicate with normal person. This huge challenge makes them uncomfortable and they feel discriminated in society. Because of miscommunication Deaf and Dumb people feel not to communicate and hence they never able to express their feelings. HGRVC (Hand Gesture Recognition and Voice Conversion) system localizes and track the hand gestures of the dumb and deaf people in order to maintain a communication channel with the other people. The detection of hand gestures can be done using web camera. The pictures are then converted into standard size with the help of pre-processing. The aim of this project is to develop a system that can convert the hand gestures into text. The focus of this project is to use deep learning technique to be applied over a dataset of static hand gesture images and infer appropriate result and later to be converted into text. The detection involves observation of hand gestures. The method gives output in text format that helps to reduce the communication gap between deaf-mute and people

2. RELATED WORK

Development of Neural Networks have enabled the research community as well as developers to provide solutions for N number of problems in the society. Especially, Convolutional Neural Networks has inspired to solve the problems which needs machines to visualize i.e. Images or Videos. Taking this unique feature of CNN into consideration many enthusiast developers have come with solution for translating sign language into text, also names as action recognition. For example, the work of, Vivek Bheda andN.DiannaRadpour mentioned in paper "Using Deep Convolutional Networks for Gesture Recognitionin AmericanSignLanguage" [1], Brandon Garcia and Sigberto Alarcon Viesca mentioned their efforts in paper "Real-time American Sign Language Recognition with Convolutional Neural Networks" [2], and there are many more. We are taking forward their work, jumping from desktop application to android application, integrating the trained CNN model into android application with highest accuracy possible.

3. METHODOLOGY

A. Data:

We use data set from GitHub repository Evil Ports [3] and also data set created of our own. The dataset contains camera snaps of static signs (hand gesture) such as alphabets A-Z. The overall dataset consists of 900 to 1000 images of each letter, which leads to huge collection and

efficiency of results. The images of signs are clicked in various background environments so as to increase the inference rate in different environments at run time. The dataset is splits into training and testing dataset at the time model training.

B. Convolutional Neural Net:

CNNs are feature extraction models in deep learning that recently have proven to be to be very successful at image recognition. As of now, the models are in use by various industry leaders like Google, Facebook and Amazon. And recently, researchers at Google applied CNNs on video data. CNNs are inspired by the visual cortex of the human brain. The artificial neurons in a CNN will connect to a local region of the visual field, called a receptive field. This is accomplished by performing discrete convolutions on the image with filter values as trainable weights. Multiple filters are applied for each channel and together with the activation functions of the neurons, they form feature maps. This is followed by a pooling scheme, where only the interesting information of the feature maps are pooled together. These techniques are performed in multiple layers. In our work, we are using MobileNet Convolutional Neural Network, TensorFlow library to work with operations related to training or retraining of MobileNet and TensorFlow Mobile Lite to run trained model on Android device.

- **MobileNet**: MobileNets are small, low-latency, low-power models parameterized to meet the resource constraints of a variety of use cases. They can be built upon for classification, detection, embedding and segmentation similar to how other popular large scale models, such as Inception, are used. MobileNets can be run efficiently on mobile devices with TensorFlow Mobile [4].
- **TensorFlow**: TensorFlow is an open source software library for high performance numerical computation. Its flexible architecture allows easy deployment of computation across a variety of platforms (CPUs, GPUs, TPUs), and from desktops to clusters of servers to mobile and edge devices. Originally developed by researchers and engineers from the Google Brain team within Google's AI organization, it comes with strong support for machine learning and deep learning and the flexible numerical computation core is used across many other scientific domains [5].
- **TensorFlow Mobile Lite**: TensorFlow Lite is the official solution for running machine learning models on mobile and embedded devices. It enables on-device machine learning inference with low latency and a small binary size on Android, iOS, and other operating systems [6].

C. Proposed Architecture:

The proposed system architecture has been shown in fig.1. This system will be built for the normal people those which will be communicating with deaf & dumb community. Users, need to hold their android phone's camera pointing towards the disabled person and start the video. The trained model at the backend will process the frames of video and recognize the hand gesture made by the disabled person. Further the semantic of signs will be converted to Speech/Voice. The key part of the system is .pb file i.e., Protocol Buffers which contains the graph of the trained model in a serialized format.

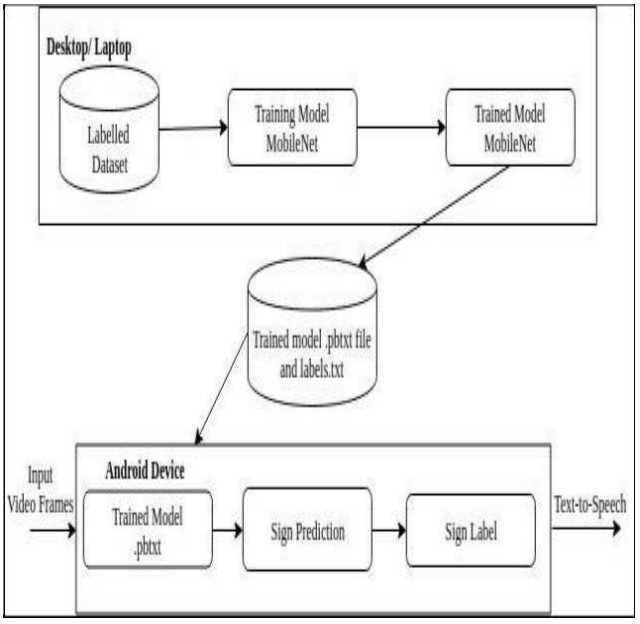

Fig 1. System Architecture

The proposed system architecture has been shown in fig.1. This system will be built for the normal people those which will be communicating with deaf & dumb community. Users, need to hold their android phone's camera pointing towards the disabled person and start the video. The trained model at the backend will process the frames of video and recognize the hand gesture made by the disabled person. Further the semantic of signs will be converted to Speech/Voice. The key part of the system is .pb file i.e., Protocol Buffers which contains the graph of the trained model in a serialized format.

4. .RESULTS

After customizing the MobileNet Convolutional Neural Network for static hand gestures (alphabets), following were the results and accuracy:

Fig 2. Alphabet "A"

Fig 3. Alphabet "B"

Fig 3. Alphabet "C"

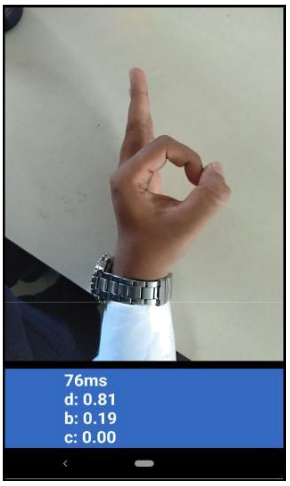

Fig 4. Alphabet "D"

5. CONCLUSION AND FUTURE WORK

Sign Language Translator will act as strong bridge of communication between normal people and deaf & Community. This application will lead people to instantly interpret the sign language without any need third person. This application will help society to communicate the deaf & dumb community and generate a bond between them. We are focusing that this application will help to interact with deaf and dumb person at places such as Shops, Hospitals, Police headquarters etc. Further implementation of our system will enable text and audio support with gesture recognition. Future research is to identify cues that not only the shape of the hand but also combined with facial expressions and posture for smooth interaction.

REFERENCES

[1] VivekBheda andN.DiannaRadpour. Using DeepConvolutional Networksfor Gesture RecognitioninAmericanSignLanguage.
[2] Brandon Garcia and Sigberto Alarcon Viesca. Real-time American Sign Language Recognition with Convolutional Neural Networks.
[3] https://github.com/EvilPort2.
[4] https://github.com/tensorflow/models/blob/master/research/slim/ne ts/mobilenet_v1.md.
[5] https://www.tensorflow.org/.
[6] https://www.tensorflow.org/lite/