Sign Language Detection

(sign to speech)

Semester Project

Session: 2020 – 2024

SUBMITTED TO

Dr. Talha Waheed

SUBMITTED BY

Umair Manzoor 2020-CS-24
Musawir Ahmed 2020-CS-46
Zain Hassan 2020-CS-134

Department of Computer Science
**University of Engineering and Technology
Lahore Pakistan**

Introduction:

The act of transmitting information from one location, person, or group to another is referred to as communication. A specified set of hand gestures with particular meanings used by hearing-impaired persons to communicate in daily life is referred to as a sign language. They communicate through body, facial, and hand motions since they are visual languages. SLR is a very vast topic for research where a lot of work has been done but still various things need to be addressed. The objective of this project is to develop a real-time SLR system using TensorFlow object detection API and train it using a dataset that will be created using a webcam.

Description:

For Sign Language, a real-time sign language identification system is being created. Its purpose is to make ordinary people understand the meaning of sign language. It'll provide assistance in recognition of sign language and ease of communication between ordinary people and special ones. It expedites the incorporation of artificial intelligence into commercial goods and offers a standard infrastructure for computer vision-based applications. With effective computer vision and machine learning functions, the OpenCV library Object recognition, face detection, and object identification algorithms classification of human actions, tracking camera and object movements, extracting 3D much more, including object models.

Technologies:

- Python
- TensorFlow object detection API
- OpenCV
- MATLAB
- Text-to-speech API

Methodology:

A labeled map is created which is a representation of all the objects within the model, i.e., it contains the label of each sign (alphabet) along with their id. The label map contains 26 labels, each one representing an alphabet. Each label has been assigned a unique id ranging from 1 to 26. This will be used as a reference to look up the class name. TF records of the training data and the testing data are then created using generate_tfrecord which is used to train the TensorFlow object detection API. TF record is the binary storage format of TensorFlow. Binary files usage for storage of the data significantly impacts the performance of the import pipeline consequently, the training time of the model. It takes less space on a disk, copies fast, and can efficiently be read from the disk.

During the training, the model has some losses as classification loss, regularization loss, and localization loss. The localization loss is mismatched between the predicted bounding box correction and the true values. The formula of the localization loss. The real-time detection is done using OpenCV and webcam again. For, real-time detection, cv2, and NumPy dependencies are used. The system detects signs in real-time and translates what each gesture means into English. From here then we can use any publicly available Text-to-speech API to get this text converted into a speech for the understanding of ordinary people.