

Project Title: Sign Language to Text

TEAM NO.: 106

NAMES OF THE STUDENTS PARTICIPATED IN THE TEAM:

SANCHIT AGARKAR, PRANAV UNKULE, PRATIK SAURKAR, CHATAK SHINDE

COLLEGE: MIT ACADEMY OF ENGINEERING, ALANDI, PUNE

SEMESTER: 7

DEPARTMENT: ELECTRONICS AND TELECOMMUNICATION

CITY: PUNE

STATE: MAHARASHTRA

PROJECT MENTOR NAME: MS. MALA MISHRA

Project Details:

An interface that bridge the communication barrier by converting sign language gesture to written text and display the same also convert the displayed text into audio.

Problem Statement:

To develop a system to convert Indian sign language into text to help hearing impaired /deaf people while communicating with the rest of the world.

Need of Project:

Sign language is a channel of conversation used by people with impaired hearing and speech. In contrast, non-signers find it utterly difficult to understand and respond to the actions, and gestures of hearing and speech-impaired people, hence qualified sign language interpreters are needed during legal and medical appointments. But, India has a total of 250 certified interpreters and so there was a need for a system that will incorporate conversation of sign language to remove this communication barrier.

There are around 300 sign languages used worldwide which include American Sign Language (ASL), Chinese Sign Language (CSL), Indian Sign Language (ISL), etc. Many kinds of research and project works are done on sign language. It is observed that most of the work is done in American Sign Language (ASL). But there is no such advancement in other sign languages. In India, deaf and dumb people use Indian Sign Language (ISL) to communicate. According to the World Health Organization, over 5% of the world population suffers from disabling hearing loss out of which 34 million are teenagers. One improvement that arrived was a requirement for educational facilities such as specialized schools for deaf and dumb people.

After these observations and statistics, it is proposed to give a platform for deaf and dumb people to express their thoughts, emotions, feelings, creativity, and explore the world. Also, by converting detected text to audio format it is assured to make communication more comfortable for such special people.

Proposed Solution:

The proposed model consists of Convolution Neural Network based Alexnet Architecture which predict the gesture into text and furthermore the predicted text is converted to audio using a python library named pyttsx3.

Technology Used:

- Computer Vision (Dataset Creation and Image Feed)
- Canny Edge Filtering (Dataset Pre-processing)
- TensorFlow (Implementation of Deep Learning Model)
- Tkinter (GUI Creation)
- Pyttsx3 (Audio Conversion)

Project Outcomes:

The system classifies the shown gesture into text on the screen and along with it gives the audio output for the corresponding gesture.

Modelling:

The project consists of main 3 steps:

1) Dataset Creation:

To create datasets and perform pre-processing operations on them, "OpenCV" is used. The dataset consists of pre-processed images taken from video frames which were captured using a laptop camera. For pre-processing, the Canny Edge Filtering technique is used.

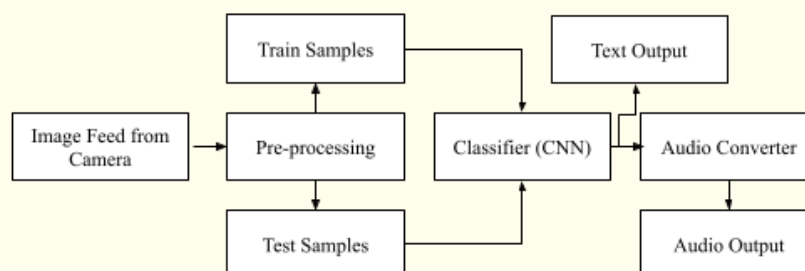
2) Training the model over dataset:ss

The created dataset is then feed to CNN model for training and the weights are saved into the Jason format which are then used for prediction purpose.

3) Real-Time Prediction and Audio Conversion:

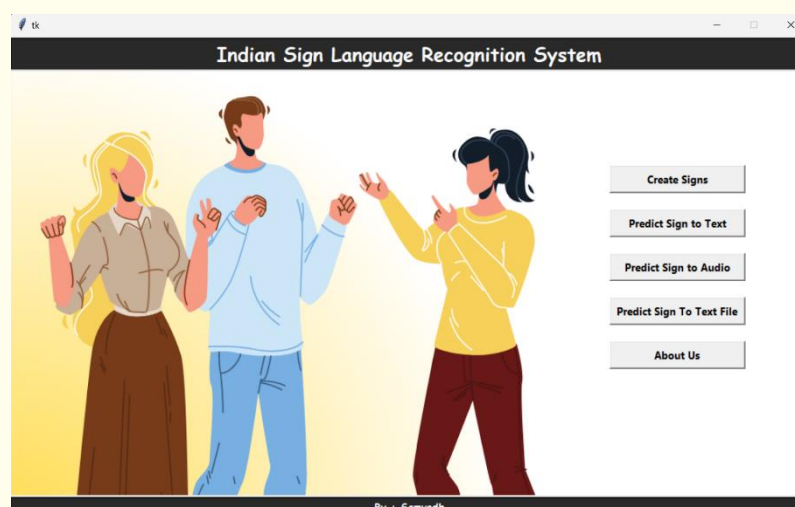
In real-time prediction the image feed from the camera is pre-processed and the pre-processed image is passed down to the saved model from classification purpose and the text is displayed over the screen, also the text is converted into audio with the help of "pyttsx3" library.

The steps can we visualized by seeing the below block diagram

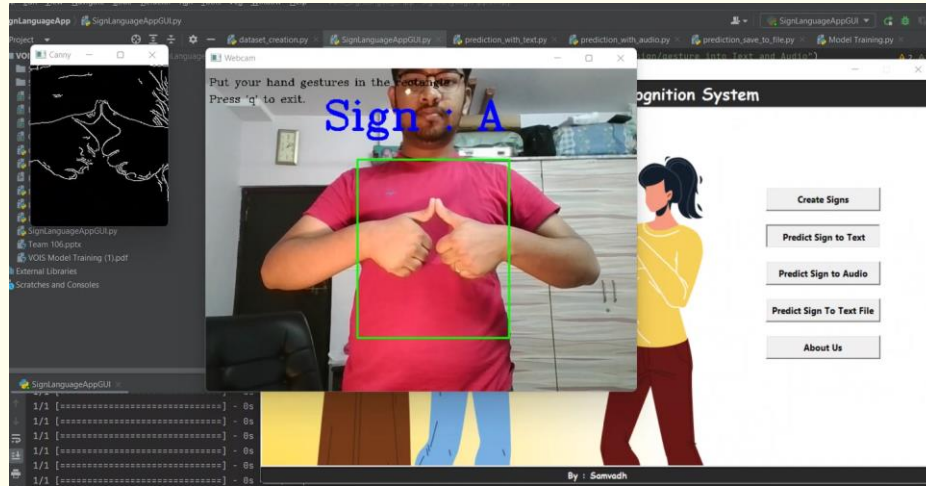


Results:

1. Graphical User Interface (GUI):



2. Real-Time Prediction Output:



Future scope for project enhancement:

The proposed system was trained on static gesture images and it can be extended for the prediction of dynamic gestures. Currently, the model is trained on 22 gestures but ISL is vast and contains thousands of gestures so this model can be extended to classify more gestures. Also, this model can only be used in certain situations due to limitations of background, so further research and work are required to remove these limitations.