

Sign Language Recognition

Context

World Health Organization's (WHO) survey states that above 6% of the world's population is suffering from hearing impairment. In March 2018, the number of people with this disability is 466 million, and it is expected to be 900 million by 2050. Also, the 2011 census of India states that 7 million Indians are suffering from hearing and speech impairment. They prefer to communicate in sign language because they can express their emotions and feelings through signs only. Hand gestures are one of the nonverbal communication modalities used in sign language. It is most often used by deaf and/or dumb individuals who have hearing or speech impairments to communicate with other deaf and/or dumb people or with normal people. Many manufacturers across the world have created various sign language systems; however they are neither adaptable nor cost-effective for end users. The task is to develop a Computer Vision algorithm which can be implemented either as a Software API or on edge platform which can recognize signs from images or videos.

Case

Physically impaired people find it hard to communicate with other humans or latest computing innovations in conversational AI. They normally communicate via sign languages which is often a mix of hand gestures and facial expressions. It would be great if there is a computer interface which can interpret their hand gestures from a single camera view and convert it into text or speech transcript. We are dealing with static signs initially for simplicity but the similar concept can be utilized for developing dynamic sign recognition. This system can help in effective human computer interaction for impaired persons which can be utilized for Customer Support, virtual meetings etc.

Problem Statement

Given a dataset of images with sign language for each character and digits, the task is to develop a Machine Learning/ Deep Learning algorithm which can automatically classify sign language from a video taken from a webcam and convert it to text/speech transcript. The goal is to develop a complete prototype which can assist the physically impaired people in interacting

with computers via webcam. The dataset for Indian Sign Language can be found at the following link:

<https://www.kaggle.com/datasets/prathumarikeri/american-sign-language-09az>. It contains images of hand gestures for each character and digits. You can utilize 80% of these images for training and remaining for validation and testing. The performance of the algorithm can be measured using various well known classification matrix and computational complexity of the algorithm. It would be great if participants can demonstrate the actual working on a camera interfaced with laptop. The use of any additional data apart from the dataset is not allowed. The participants can use any programming language and machine learning framework for developing their model. The values of hyperparameters used for training the algorithms, preprocessing techniques and any limitations should be clearly mentioned in the presentation.

Things to be submitted

- A video recording of end-to-end working of the algorithm
- A document / presentation. It can contain the following details:
 - Model Description
 - Pre/Post Processing Techniques
 - Choice of Hyper parameters
 - Results
 - Limitations and Future Enhancement
- Code bases (which are hosted on Github)

Good to have features

The participants can also try to deploy their algorithm which can help in development of the actual system. The deployment can be on:

1. Embedded development boards like Jetson Nano or Raspberry Pi and demonstrate the working via camera interfaced with the board.
2. Deployed as an API on any cloud platform which takes an uploaded image as an input and gives the classified sign in return.

Final Notes

The goal of the competition is not only to design a best-performing recognition algorithm but understanding the complete workflow from data preprocessing to actual deployment of the algorithm and it would be great if participants can try to complete the entire flow.