

Sign to Text

Abstract

This project will be beneficial to the deaf and dumb community, as it will increase their access to communication and services. It will also help to raise awareness of the lack of resources available to them and help to ensure that they have the same opportunities as everyone else. By creating an application that can accurately interpret sign language, we will be helping to bridge the communication gap and ensure that everyone has access to the same information.

Sign Language and its component

1. Finger Spelling

2. Word level sign vocabulary

3. Non Manual Features

Social cause

The goal of this cause is to raise awareness of the severe lack of resources available to deaf and blind people, and to ensure that they have access to the same opportunities as everyone else. This includes providing interpreters and other tools to help them communicate, as well as providing education on sign language and Braille. It also involves advocating for increased access to employment opportunities and other services. Finally, it is important to recognize the unique needs of people with hearing and vision impairments and to ensure that their voices are heard and respected.

Methodology

- Processing the dataset
- Gesture Classification
- CNN Classifier Model
- Finger Spelling Sentence formation

Gesture Classification

Algorithm layer 1

Apply gaussian blur filter and threshold to the frame taken with opencv to get the processed image after feature extraction.

2. This processed image is passed to the CNN model for prediction and if a letter is detected for more than 50 frames then the letter is printed and taken into consideration for forming the word.
3. Space between the words are considered using the blank symbol.

Algorithm Layer 2

- We detect various sets of symbols which show similar results on getting detected.
- We then classify between those sets using classifiers made for those sets only.
- In our testing we found that following symbols were not showing properly and were giving other symbols also:
 1. For D: R and U
 2. For U: D and R
 3. For I: T, D, K and I
 4. For S: M and N

Convolutional Neural Networks

Convolutional Neural Networks are a type of Artificial Intelligence that helps computers recognize objects and shapes in pictures. It works by taking pictures and breaking them down into lots of tiny pieces, then looking at each piece to figure out what the picture is.

This is like a puzzle, but for a computer. It helps computers become better at recognizing patterns, like people's faces, animals, and even things like roads and buildings.

CNN Classifier Model

Finger Spelling Sentence Formation

IMPLEMENTATION

1. Whenever the count of a letter detected exceeds a specific value and no other letter is close to it by a threshold we print the letter and add it to the current string(In our code we kept the value as 50 and difference threshold as 20).
2. Otherwise we clear the current dictionary which has the count of detections of present symbol to avoid the probability of a wrong letter getting predicted.
3. Whenever the count of a blank(plain background) detected exceeds a specific value and if the current buffer is empty no spaces are detected.
4. In other case it predicts the end of word by printing a space and the current gets appended to the sentence below.

Autocorrect feature

A python library **Hunspell_suggest** is used to suggest correct alternatives for each (incorrect) input word and we display a set of words matching the current word in which the user can select a word to append it to the current sentence. This helps in reducing mistakes committed in spellings and assists in predicting complex words.

Challenges Faced

- One issue was to select a filter for feature extraction. We tried various filter including binary threshold, canny edge detection, gaussian blur etc., of which gaussian blur filter was giving better results.
- Issues were faced relating to the accuracy of the model we trained in earlier phases which we eventually improved by increasing the input image size and also by improving the dataset.

Software Requirements

- Python 3.6.6
- Tensorflow 1.11.0
- OpenCV 3.4.3.18
- NumPy 1.15.3
- Matplotlib 3.0.0
- Hunspell 2.0.2
- Keras 2.2.1
- PIL 5.3.0

Limitations of our model

- The model works well only in good lighting conditions.
- Plain background is needed for the model to detect with accuracy.

Conclusion

- In this report, a functional real time vision based american sign language recognition for D&M people have been developed for asl alphabets.
- We achieved an accuracy of 98.00% on our dataset.
- Prediction has been improved after implementing two layers of algorithms in which we verify and predict symbols which are more similar to each other.

Future Scope

- ❖ We are planning to achieve higher accuracy even in case of complex backgrounds by trying out various background subtraction algorithms.
- ❖ We are also thinking of improving the preprocessing to predict gestures in low light conditions with a higher accuracy.

Mentored by

Pratham Tripathi

Parinay Rana

Efforts by

Sarika Lakhotia 21bme052

Gautam jangir 21bch048

Yashitaa Arya 22bcs120

Anuj Nautiyal 22bme024