SIGN LANGUAGE GESTURE RECOGNITION

Kaustav Vats
CSE Department
IIIT Delhi
New Delhi, India
kaustav16048@iiitd.ac.in

Lakshya Bansal
CSAM Department
IIIT Delhi
New Delhi, India
lakshya16240@iiitd.ac.in

Deepanshu Badshah
ECE Department
IIIT Delhi
New Delhi, India
deepanshu16144@iiitd.ac.in

I. PROBLEM STATEMENT

This project aims to compare and analyze the following two approaches :
1) Hand Gesture Recognition using SVM.
2) Hand Gesture Recognition using CNN.

II. LITERATURE REVIEW

Sign language recognition and gesture recognition are two major applications for hand gesture recognition technologies. The main goal of sign language recognition is to automatically interpret sign language to help the deaf and dumb people to communicate among themselves or with normal people conveniently.

The current ways, to the best of our knowledge, to perform hand gesture and sign language recognition mainly focus on pre-processing the data and applying CNN, ANN and SVM with different changes for recognition. The following are the papers which we are referring for our project:

[1] recognizes using various feature extraction techniques like shape descriptors, SIFT and HOG individually along with SVM classifier.

[2] uses CNN (max pooling strategy) with dropout to classify images of both the the letters and digits in American Sign Language.

[3] uses CNN with stochastic pooling strategy for classification of gestures in selfie videos. CNN training is performed with 3 different sample sizes, each consisting of multiple sets of subjects and viewing angles.

## III. DATASET

A. Argentinian Sign Language Dataset - 1

Dataset includes 3200 videos of 10 non-expert subjects doing 5 repetition of 64 different hand signs. Some of the common LSA hand signs were used for creating dataset. Data is divided into two sets:

- The first set contains 23 one-handed signs. Recording for this set was done in natural lighting.
- The second set contains 41 signs, 22 two-handed and 19 one-handed signs. Recording for this set was done in an indoor environment, with artificial lighting.

For both sets all subjects wore black clothes with a white background. To simplify segmentation subjects also wore fluorescent colored gloves. These simplified the dataset pre-processing by removing issues related to different skin color.

Fig. 1. Samples of extracted frames
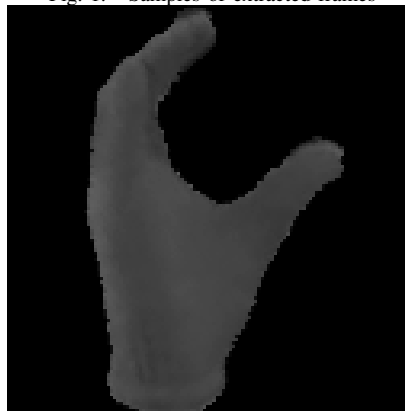
IV. PROJECT UPDATE

Work we have done till now:

1) We have collected videos from Argentinian Sign Language data set.
2) Extracted the relevant frames from each video.
3) Converted the frames into grey-scale
4) bag of words model is implemented over the extracted frames of the training videos and the histograms are computed
5) SVM is trained over the training histograms computed above and classification is done one the test histograms computed from the test videos.

V. RESULTS

Since the training data is large and the computation of the clusters in the bag of words model takes time, the results could not be obtained in time for the submission. The up-to-date code for the project update has been submitted for reference.

REFERENCES

[1] Juhi Ekbote, Mahasweta Joshi, Indian Sign Language Recognition using SVM and ANN classifiers
[2] Vivek Bheda, N. Dianna Radpour, Using Deep Convolutional Networks for Gesture Recognition in American Sign Language
[3] G.Anantha Rao, K.Syamala, P.V.V.Kishore, Deep Convolutional Neural Networks for Sign Language Recognition