CS 584 Machine Learning (Spring 2023)

Sign Language Recognition

Ronil Christian (A20489525)
rchristian@hawk.iit.edu
Illinois Institute of Technology

DESCRIPTION

According to World Health Organization, approximately 466 million people have disabling hearing loss, which represents 5% of the world's population. In the United States, the National Institute on Deafness and Other Communication Disorders (NIDCD) reports that approximately 15% of American adults, or 37.5 million people, have some degree of hearing loss. This includes 2-3 out of every 1,000 children born with hearing loss or who develop hearing loss in childhood.

The deaf and hard of hearing community faces several challenges, including communication barriers, social isolation, and discrimination. Many members of the community rely on sign language, captioning, or other assistive technologies to communicate, and they may face difficulties in accessing education, employment, and healthcare.

Efforts to address these challenges have introduced the sign language gesture classification as problem of recognizing and interpreting sign language gestures used to convey messages, which uses hand gestures, facial expressions, and body languages to convey the messages.

In my work, I'd like to address this as a classification problem that use cameras to capture sign language and machine learning algorithms to classify signs and sequences.

PAST WORK

Previous works have used linear classifiers achieving 96% accuracy on 10 classes. HMM's have also achieved good accuracies, but they require a 3-D glove that used hand movements. Neural networks have been used as well, and are significant because they learn the most important classification features. Most networks have been shallow and they require more data and time. And however, the computational complexity has increased exponentially.

PROPOSED WORK

Deep learning models, particularly convolutional neural networks (CNNs), have shown promising results in sign language recognition due to their ability to learn hierarchical features and their high performance in image and video recognition tasks.

Here, I intend to present an ASL recognition system that uses Convolutional Neural Networks (CNN) in real time to translate a video of a user's ASL signs into text. Our problem consists of three tasks to be done in real time:

1. Obtaining video of the user signing (input)
2. Classifying each frame in the video to a letter
3. Reconstructing and displaying the most likely word from classification scores (output)

PRELIMINARY PLAN (MILESTONES)

Task	Estimated completion
Literature Review	15-Mar
Data Gathering	23-Mar
Intermediate Report	31-Mar
Coding	14-Apr
Presentation	14-Apr
Final Report	30-Apr

1. Literature Review: I will be going through some of the papers which have previously tackled this problem and try to identify areas of improvements.
2. Data Gathering: Manually, I be using OpenCV to capture images for the letters A-Z. Alternatively, there are datasets available on Kaggle.
3. Coding: In this part, I intend to compare various algorithms to find out which works best for the given problem. To make it end-to-end, I will also be creating a UI for the application.
4. Report: Will be using LateX.

REFERENCES

1. Shin J, Matsuoka A, Hasan MAM, Srizon AY. American Sign Language Alphabet Recognition by Extracting Feature from Hand Pose Estimation. Sensors (Basel). 2021 Aug 31;21(17):5856. doi: 10.3390/s21175856. PMID: 34502747; PMCID: PMC8434249.
2. Garcia B., Viesca S.A. Real-time American sign language recognition with convolutional neural networks. Convolutional Neural Netw. Vis. Recognit. 2016;2:225–232.