American Sign Language Alphabets Recognition / Classification

Team: Ajay Karthick Senthil Kumar & Firdose Shaik

We propose to develop a deep learning model to classify images of the American Sign Language Alphabets. American Sign Language (ASL) is a commonly used visual language by the people who are hard of hearing in the United States. It uses hand gestures to convey thoughts instead of the spoken words. ASL has 26 signs for the alphabets of the English language.

Problem:
The hearing-impaired person can communicate to only limited number of people using ASL since very few people understand ASL. Hence developing a Machine Learning Model to recognize ASL alphabets will pave a way to create an effective communication channel. It is interesting because it tries to address a real-world problem and an attempt to resolve it via machine learning algorithms.

Datasets:

- https://www.kaggle.com/datasets/grassknoted/asl-alphabet
- Training dataset of 87000 images
- 29 classes - 26 alphabets (A-Z), Space, Delete and Nothing.

Approach:
We plan to make use of the Deep Learning algorithm of Convolutional Neural Network (CNN) for the task of classification of the ASL Alphabets. First, we do the image data pre-processing then we build the Deep Convolutional Neural Network, compile and train the model with the training dataset, finetune the model, evaluate the model with the test dataset and then use it to recognize the real-time input (ASL alphabet) given by the user.

Key components of our approach:
As we deal with the unstructured data in our model, it is very important to preprocess the images data. Data Preprocessing phase may include steps like resizing all images data to uniform size, normalization and image augmentation. Once we preprocess the data, we build the CNN model from scratch rather than using pre-trained model (Transfer Learning). We intend to build the CNN with deep convolutional layers, max pooling layers, ReLU activation, SoftMax activation for classification. We will also have dropout layer to prevent the model to overfit. We intend to use Adam Optimization Algorithm but also try out different Optimization Algorithms that yields best accuracy. Hyperparameter tuning will help to maximize our final model's accuracy. We are planning to capture real-time ASL alphabet hand gestures shown by the end user feed it to the trained model, and the model will classify and gives the corresponding English Alphabet.

While most of the competing approach uses Transfer Learning where it uses pre-trained models to achieve this task, we intend to build the entire CNN model from scratch and train it with the sufficiently large training data. If time permits, we are also looking to work on generating the entire word/sentence formed by the ASL alphabets hand gestures shown by the user in addition to classify the alphabets.