Project Proposal

Topic / Project Name
American Sign Language (ASL) Detection using Convolutional Neural Networks

Group Members
Nikunj Lad – 001422467 (Group 2)

Introduction
1. **Overview of the project:**
   We are privileged as humans to talk and converse and communicate with each other. But what about those people who can't talk or speak? Such people find it hard to convey their ideas. In view of that, what if a system exists which can help in identifying hand signs of people and can categorize them from live stream of video like a webcam or a camera. This project intends to make use of Convolutional Neural Networks to identify and classify hand signs into corresponding classes – like say letter A, B, etc.

2. **Objectives for this project:**
   Primary objective of this project is to solve a real-world communication problem between humans who can speak and those who cannot speak. Another major objective is to exploit the power of deep learning and high-performance computing platforms to train a massive dataset of 78000 images of hand signs and develop a model to help identify such images on the fly.

Methodology
For this project, I intend to use Inception-Net and VGG Net along with some custom architecture to develop a robust classifier. Also, I will be using PyTorch as a library for programming since it will give me finer control over data-loaders and network components. Also, PyTorch has a very solid data-parallelism pipeline for training models on CUDA and also for multi-GPU execution for distributed training.

Specifications and Description of Dataset
The dataset if an American Sign Language dataset, having hand sign of each letter in the English vocabulary (A-Z) and some other custom hand signs to denote Space, Delete and Blank representations. Each category has 3000 images, since we have 26 alphabets and an extra 3 custom categories, a total of 29x3000 = **87000** images is there. This is a **1.27GB** large dataset.

Data Sources
The dataset was taken from this Kaggle - https://www.kaggle.com/grassknoted/asl-alphabet

Additional Information
Since this dataset is pretty huge and considering the limited availability of GPU nodes on the Discovery cluster, I will be using a GCP instance which I have available with me with 4 Nvidia Tesla T4 GPU's for my training.