Transfer Learning for American Sign Language Dataset

Josh Hills - 101142996
COMP 4107
Professor Holden

February 28, 2023

Introduction

I've worked on lots of neural network problems in the last 2 years, starting in COMP3106 (Intro to Artificial Intelligence) I began experimenting with neural networks and since have gained a way better grasp on what is going on. One domain of neural networks that I have never explored is transfer learning, my goal is to understand transfer learning deeply by the end of this project.
I will be working on a computer vision problem, using a limited dataset of images of hand signs, classify the image as the corresponding letter in ASL.

Methods

As mentioned, transfer learning will be used for this. What is transfer learning? Transfer learning is a machine learning technique where a model trained on one task is re-purposed and adapted to perform a different but related task. In transfer learning, a pre-trained model is used as a starting point for a new task, rather than training a new model from scratch. This requires a trained model in similar scope to the problem at hand, for this I will be using a pre-trained classifier based on the HaGRID (Hand Gesture Recognition Image Dataset). I will then have to fine-tune the model for the new task, in this case classifying the letter of the corresponding hand sign.

Datasets

Two datasets are technically used for this project:

1. The HaGRID dataset, I won't actually be training a model because this dataset is huge, 716GB to be clear. This consists of 552,992 (1920x1080) images of people standing variable distances from their cameras, in variable lighting divided into 18 classes of gestures (ex. dislike, fist, okay, etc.).

2. For the ASL alphabet dataset, there are many to chose from. I'm excited to explore how variable resolutions of images, sizes of datasets will affect performance. But there is a baseline dataset called Sign Language MNIST, found on kaggle. The thing I like most about this dataset is that it's small, so I expect transfer learning to help a lot here.

Proposed validation/analysis strategy

As previously mentioned, my goal is to investigate transfer learning, so I will be comparing the performance of my pre-trained model with and without transfer learning. The validation strategy is going to consist of a simple accuracy metric. I will split the data into 80/20 training/testing sets. The model will be evaluated on its ability to correctly classify images in the testing set. I'll also make a confusion matrix to visualize the accuracy.

Weekly Schedule

I have 5 weeks before the report is due:

1. I plan on delving into the data, getting a grasp on how the HaGRID model is trained and exploring deeper the ASL datasets. Also get the HaGRID model working locally

2. Create a model for the ASL dataset without using transfer learning

3. Train a model using transfer learning

4. Compare models performances, create visualizations (graphs+confusion matrix)

5. Write report