Leaksmy Heng

CS5330

March 26, 2025

Final Project Design

Final Project

I would like to develop a real-time American Sign Language (ASL) recognition system that could interpreting static ASL alphabet gestures using deep learning. This will allow individuals who are deaf or have difficulties hearing communicate more efficiently and effectively with others. The dataset could be obtained from Kaggle in ASL Alphabet. This dataset contains 87,000 images which are 200x200 pixels. I will use ResNet-50 to train the model, and if I have more time, I'll also going to use EfficientNet than compare it to see which one perform better.

Deep network architecture:

- Using ResNet-50 (and Efficient net if time permit)
- Real-time integration using TensorFlow Object Detection API

Design solutions:

- Use pre-train model ResNet-50 (and Efficient net if time permit)
- Split dataset into 70% training, 15% validation and 15% testing
- Optimize hyperparameters with learning rate, batch size, etc.
- Integrate this with real-time system (my camera)
- Evaluate the model using static gesture recognition through F1-score, accuracy, precision and recall.
- Latency (frame per second) to ensure real-time performance (if time permits)

For computing power, I am planning on requesting the GPU usage from Northeastern and/or using the GPU provides by Kaggle (NVIDIA TESLA P100 GPU).

Good outcome of the projects mean that my accuracy is at 80% or more on static ASL alphabet gesture and the real time performance is not lagging that much.

References

https://www.kaggle.com/datasets/grassknoted/asl-alphabet?select=asl_alphabet_train

https://www.kaggle.com/docs/efficient-gpu-usage

https://www.youtube.com/watch?v=pDXdlXlaCco