EECS442 Project Proposal

Serena Su Wang, Hogan Gk Mastanduno, Yajwin Manojjkumar Jain, Aman Jalan

TOTAL POINTS

100 / 100

QUESTION 1

1 No need to mark anything 100 / 100

✓ **+ 100 pts** **Non-custom proposal**

+ 100 pts Custom proposal

You did a great job! The plan looks pretty good to me and I think you should be all good if you can follow that.

Here are a few suggestions:
- In terms of the data, if you find one dataset might not work well, you can try a different one. Using a subset of a dataset could largely save training time and let you if your model works or not.
- Great to see you have those estimated numbers (dataset size, # of images)!

Project Title: Sequential Sign Language Recognition & Transcription

Group Members: Aman Jalan (ajalan), Serena Wang (wserena), Yajwin Jain (yajwin), Hogan Mastanduno (hoganmas)

Group Dynamics: We will communicate mainly through an iMessage group, where we expect to get back to each other within a day. We plan on meeting online and in-person on Tuesdays, Thursdays, and Fridays. We'll resolve group issues in a democratic way, and will be sure to be open to positive feedback from each other.

Problem Statement: Each sign language gesture represents an alphabet, and a sequence of gestures represents a word. When we try predicting words from single gesture models, we lose a lot of accuracy because many gestures are very similar in sign language. A better way to predict sign language translations is to look at sequences of gestures, as that gives us extra information such as co-articulation and spontaneous sign production. We plan to build a model that uses sequential sign language gestures to accurately predict their translations, and make communication in the real world easier.

Approach: We plan to train a neural network to the ChicagoFSWild fingerspelling dataset. We will use this neural network to translate images of ASL handshapes into text. We will compare the efficacy of this approach to a baseline nearest-neighbor search. We also plan to use transfer learning and use a separate dataset to get more accurate results.

Data: We plan to use the ChicagoFSWild: American Sign Language fingerspelling dataset. This data set contains 7304 ASL fingerspelling sequences signed by 160 signers.

Computational Resources: We will likely need a considerable strong GPU to train our neural network on the entire dataset (~14GB). One of our team members has a computer with a NVIDIA GTX 1070, which is decently strong enough. However, should we need something stronger, we will consider using the computer labs in the Duderstadt.

Evaluation: We can measure quantitative metrics such as recall, accuracy, and F1-scores. We can compare output results with some baseline results to get these metrics. F1-scores are a great evaluation metric, as they help catch models that return either too many false positives or too many false negative scores. Thus, this will help us to build and evaluate a balanced model. Another good form of evaluation is just by checking the readability of translations, and this should be dependent on the end-user feedback of the system.

1 No need to mark anything **100 / 100**

✓ **+ 100 pts** **Non-custom proposal**

+ 100 pts Custom proposal

You did a great job! The plan looks pretty good to me and I think you should be all good if you can follow that.

Here are a few suggestions:
- In terms of the data, if you find one dataset might not work well, you can try a different one. Using a subset of a dataset could largely save training time and let you if your model works or not.
- Great to see you have those estimated numbers (dataset size, # of images)!