2021 CSE 40625 Final Project Status Report I

Prepared by

*Carter Goldman, Livia Johan, Christina Youn*

University of Notre Dame,

Notre Dame, IN, 46556

**1. Project Goal and Milestone**

For our project, we plan to create an image recognition model that can recognize American Sign Language letters (excluding J and Z because they require motion). For the first milestone, we said that we would finish preprocessing all of the images in [this](https://www.kaggle.com/datamunge/sign-language-mnist) dataset by making sure that the images are black and white, a uniform size, and cropped correctly. We also said that we would start creating a simple, baseline machine learning model.

**2. Tasks Completed So Far**

The dataset that we chose was pre-split into training and testing datasets (approximately 80:20 split). We first loaded the dataset into pandas dataframes. For the labels, we used one-hot encoding because we were dealing with categorical data.

Next, we would have done image preprocessing. However, we were lucky that our dataset was already quite clean, which meant that we didn’t have to do much. Each image was already black and white, 28 x 28 pixels, and cropped to the hand. Here are nine example input images:



After getting these images, we created a simple convolutional neural network model. Our current model is quick to overfit to the training data, but that’s just because we only crafted our initial model to verify that our data will be compatible with the training. The initial model has around 82% accuracy on the test data.

Our progress is aligned with the planned milestone. We will now work on improving our initial model, testing the model on our own photos, and taking input from a live video stream.

**3. Team contribution**

How each of the team members contributed to the project so far.

Due to time conflicts, we were not able to meet as a group of 3.

Carter & Livia:

* Load in data from Google Drive
* Process the data and format it properly so it can be used in the CNN
* Plot images
* Create initial model and test script
  + We were unable to get the shape of the dataset correct to fit it into the model

Christina:

* Optimized and organized the code
* Debugged issues with the shape of the dataset
* Got the model to run successfully
* Drafted the Status Report