<div align="center">

CS4180 Deep Learning
Project proposal: Fingerspelling to text translation using RNN's

</div>

Students:

Harleman, Bram	4361105
Van Wijk, Robert	4313968
Klein, Keith	4298292
Ge, Zhouxin	4272358

Motivation

Sign language is a very important means of communication for hearing impaired people. Although Fingerspelling is not commonly used in casual conversation, it is still used to convey information such as the names of places and people. For our project, we want to investigate whether the use of a Recurrent Neural network(RNN) can significantly improve the performance of a Fingerspelling translator. This leads to the following research question:

<div align="center">

"**How does model architecture affect the performance of sign language word classification from a sequence of images depicting the words' letters?"**

</div>

Dataset

https://www.kaggle.com/datamunge/sign-language-mnist
For a base dataset, we will use the Sign language MNIST dataset from the Kaggle website. This dataset contains labeled training and test data for the sign language alphabet. with the exception of J and Z. Since these letters require a gesturing motion. Each letter is represented as a 28x28 grayscale image with pixel values between 0 and 255. The dataset contains 27455 cases of training data and 7172 cases of test data. Using sequences of this data, we will create common words and names that will be used to train the RNN.

Approach

We would like to try out different architectures for a recurrent neural network and come up with some performance metric (either literature or our own). We expect the following results from these experiments: An ordered list of best-performing architectures that we tried. *How we want to evaluate the results is something we still have to figure out*.

Feasibility

We think that this research is feasible because it is not an overly complex network to create and a lot of research in this field has already been performed where we could learn from. Also there are numerous models for image classification and sequence transduction publically available.

Timeline

The timeline of next quarter will look like this:

Task	Person	W1	W2	W3	W4	W5	W6	W7	W8	W9
Introduction to coding	All		█	█						
Setting up simple classifier	Keith, Robert				█					
Dataset retrieval	Bram			█						
Creating sequence list	Zhouxin				█					
Milestone report	All					█				
Training RNN	Keith, Robert						█			
Optimize RNN	Bram						█	█		
Retrain RNN	Zhouxin							█		
Test RNN	Keith, Robert								█	
Final report	All									█

Literature

A first look at the literature resulted in the following three references that can provide a starting point for this research:

[1] A. S. Nikam and A. G. Ambekar, "Sign language recognition using image based hand gesture recognition techniques," (https://ieeexplore.ieee.org/document/7916786)

[2] Vivek Bheda and Dianna Radpour, "Using Deep Convolutional Networks for Gesture Recognition in American Sign Language" (https://arxiv.org/ftp/arxiv/papers/1710/1710.06836.pdf)

[3] Maraqa et al. (2012) "Recognition of Arabic Sign Language (ArSL) Using Recurrent Neural Networks" (https://file.scirp.org/pdf/JILSA20120100003_97214394.pdf)