CS6140 Final Project Plan

Group members:
Haotian Shen, Jake Van Meter, Yihan Xu, Qiaozhi Liu

Project description:
Our team is working on an exciting project that aims to bridge the language barrier between English and Spanish speakers by developing a system that can automatically translate English handwritten or typed text into Spanish. This system has potential use cases in a variety of settings, such as education, business, and government, where individuals need to understand documents or texts written in a foreign language.

The task of translating handwritten or typed English text to Spanish involves two distinct steps. First, we need to develop a model that can accurately recognize English text from an image, whether it's handwritten or typed. To achieve this, our team plans to train a convolutional neural network (CNN) using learned embeddings based on a transformer architecture. The use of a transformer-based architecture will allow our model to effectively capture the complex spatial relationships between the pixels in the image, leading to more accurate text recognition.

Once we have accurately transcribed the English text from the image, the next step is to translate it into Spanish. To achieve this, we plan to use three different machine learning methods to generate machine-based translations of the English text. These methods will be carefully chosen based on their performance in previous studies and their ability to accurately capture the nuances of the English language.

To evaluate the accuracy of our system, we will use Google Translate to manually translate the ground-truth English text from the image into Spanish. We will then use these translations as our ground-truth Spanish text for comparison against the output generated by our machine translation models. This evaluation process will allow us to quantify the performance of our system accurately and identify areas for improvement.

Dataset description:
Dataset used for handwritten text recoginition: https://paperswithcode.com/dataset/iam
Dataset used for translation: text translation: Tab-delimited Bilingual Sentence Pairs from the Tatoeba Project (Good for Anki and Similar Flashcard Applications) (manythings.org)

Reference:
1. Easter2.0: Improving convolutional models for handwritten text recognition | Papers With Code
2. NEURAL MACHINE TRANSLATION BY JOINTLY LEARNING TO ALIGN AND TRANSLATE 1409.0473.pdf (arxiv.org)
3. Attention Is All You Need [1706.03762] Attention Is All You Need (arxiv.org)
4. Book: Deep Learning with Python, Second Edition

ML solution:
1. Extract English text from images: CNN based on transformer embeddings.
    a. Our system uses advanced neural network techniques to recognize English text from images, including 2D Convolution, Batch Normalization, ReLU, Dropout, Dense Residual connection, and Squeeze-and-Excitation module. The techniques extract features, normalize outputs, introduce non-linearity, prevent overfitting, enable better information flow, focus attention on important features, and use CTC loss for training on sequence prediction tasks.
2. Translate English text to Spanish
    a. Three methods for English-to-Spanish translation include: 1) Custom deep learning model with attention mechanism, which requires extensive training on large datasets to achieve high translation accuracy. 2) Fine-tune pre-trained Seq2Seq transformer, adapting a pre-trained model to our specific task to achieve higher accuracy with less data. 3) Fine-tune pre-trained BERT transformer model, leveraging its pre-trained language representations to achieve high accuracy with minimal data.