**Sign Sight**

**Group Members**: Abbey Liu, Dalvir Singh

**Topic**: Computer Vision, Multi-class Classification

**Motivation**:

* Sign language is defined as a complete, natural language that uses movements of the hands and face to communicate. It contains a grammatical structure and has the same linguistic properties as spoken languages. The type of sign language used differs by region and country, though they may share certain features or gestures. Sign language is used often by those who are mute, deaf, or hard of hearing.
* While there are communication barriers between spoken languages, an even larger barrier exists between signed and spoken languages. Most daily activities, such as asking for help at a store and ordering a meal, assume that a person is able to use spoken language. Thus, it is common for sign language speakers to become a segregated minority in numerous places due to the language barrier. A more consequential example of this barrier exists in the healthcare domain. When the signed/spoken communication barrier is combined with differences in health literacy, the communication barrier becomes even more prevalent. Overall, the lack of general awareness of sign language combined with a shortage of interpreters further propagates the segregation and alienation of sign language users in a dominantly spoken language society. In this project, we aim to create a machine learning model that can recognize American sign language (ASL) through images to help dismantle the communication barrier. We plan to deploy it with a web application, so that it can be accessible to a broader audience.

**Project Vision**:

* *Ideal Project Results* : A web application that can recognize ASL in real time through video input
* *Reality Results:* A web application that users can upload images and to recognize the ASL.

**Data Sets**:

* [American Sign Language Dataset](https://www.kaggle.com/ayuraj/asl-dataset)
  + 2,515 PNG images of size 400x400
  + Images are the ASL signs for all 26 letters and 10 digits
    - 65-70 images for each gesture
* [Sign Language MNIST](https://www.kaggle.com/datamunge/sign-language-mnist)
  + 27,455 training and 7172 testing grayscale images of size 28x28, saved pixel values in CSV files
  + Images are the ASL signs for all 26 letters and 10 digits
  + Dataset was created to resemble the classic MNIST handwriting dataset

**Our Approach**:

* To achieve our project vision, we plan to utilize the CNN architecture which is a popular method for image recognition. CNN is a neural network commonly used in image recognition tasks which extract features(filters) from images and create learnable parameters to perform the classification. We will train and test a Keras CNN model on our dataset and integrate the best performing model into a React web application.

**Related Literature**:

* [Real-Time American Sign Language Recognition from Video Using Hidden Markov Models](https://www.aaai.org/Papers/Symposia/Fall/1996/FS-96-05/FS96-05-017.pdf)
  + *Description*: In this paper, the researchers created a real time system that interprets American Sign Language using Hidden Markov Models. The researchers focused on gesture signing compared to finger signing because it is more popular and follows the pace of spoken conversations. The system tracks hands by color(glove or skin) and tracks their shape, orientation, and gesture which serves as the input to the HMM model to recognize the signed gesture.
  + *Technique*: Hidden Markov Models
  + *Results*:
    - The research conducted two experiments:
      * The first experiment tracked hands wearing colored gloves and achieved an accuracy of 99%
      * In the second experiment, they used hands without gloves and achieved a 92% accuracy
* [Sign Language Recognition Using Convolutional Neural Networks](https://link.springer.com/content/pdf/10.1007/978-3-319-16178-5_40.pdf)
  + *Description*: In this paper, the researchers aimed to build a real time system to recognize 20 Italian gestures using Microsoft Kinect, CNN, and GPU acceleration. The CNN model was used to extract features from the frames and an artificial neural network used for the classification. Their system performed well regardless of user or surrounding.
  + *Technique*: CNN(Feature Extraction) & ANN(Classification)
  + *Results*:
    - The study observed a validation accuracy of 91.70% (8.30% error rate) for their best model.

**Tasks (Each task will be shared evenly between the group members)**:

* Gather image data, and perform exploratory data analysis (EDA)
* Build and train various machine learning models
* Design and develop a web application
* Integrate the model with the web application
* Deploy the web application

**Timetable**:

* Week 1
  + Gather image data of sign language hand gestures
  + Perform EDA
  + Clean any data, if necessary
* Weeks 2-4
  + Implement various models and train them on the image data. Possible models to explore include:
    - CNN
    - HMM
  + Choose one (or more) well-performing models to integrate in the web application
* Weeks 5-6
  + Create, integrate and deploy web application

**Responsibilities**:

Abbey and I both want experience working on each part of the project so we have agreed to share the responsibility for all the parts.

**Non-literature References (used in Motivation section)**

* [National Institute of Deafness and Other Communication Disorders](https://www.nidcd.nih.gov/health/american-sign-language)
* [Human Rights Watch](https://www.hrw.org/news/2018/09/23/sign-language-key-deaf-peoples-rights#:~:text=Lack%20of%20awareness%20of%20sign,services%20mandated%20to%20assist%20them)
* [Disability Studies Quarterly](https://dsq-sds.org/article/view/316/380#:~:text=It%20is%20more%20common%20for,deaf%20youths%20integration%20into%20society.)
* [Boston University article](https://www.bu.edu/sph/news/articles/2018/healthcare-language-barriers-affect-deaf-people-too/)