

Assignment 4 Part 1: Data Collection of ASL Images

Deadline: Monday October 7, 2019 at 9:00pm

Late Penalty: There is a penalty-free grace period of one hour past the deadline. Any pictures that are submitted between 1 hour and 24 hours past the deadline will receive a 20% grade deduction, and an additional 20% per 24 hours after that.

One of the main reasons for the success of deep learning is the availability of very large datasets. Where do these datasets come from? Sometimes it involves hard work! In the tutorial on Thursday October 3rd at 10am, all students in the class will create data for the dataset to be used in Assignment 4. The purpose of this assignment is to help you gain experience gathering your own data set, and understand some aspects of a data cleaning process.

1 American Sign Language

American Sign Language (ASL) is a complete, complex language that employs signs made by moving the hands combined with facial expressions and postures of the body. It is the primary language of many North Americans who are deaf and is one of several communication options used by people who are deaf or hard-of-hearing.

The hand gestures representing only the English alphabet are shown below. The goal in this assignment (in Part 2) is to classify a subset of these hand gesture images using Convolutional Neural Networks (CNNs). We will focus on signs that do not require movement, specifically, the hand signs for one of the letters A-I, and K. In part 2 you will build a classifier which determines which letter is being represented.

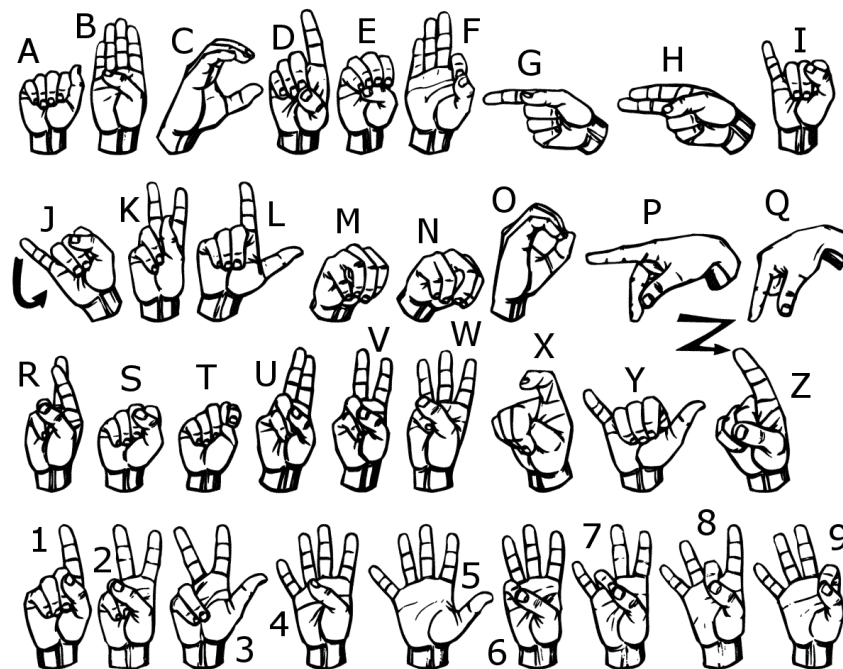


Figure 1: ASL Letter Signs taken from https://qualityansweringservice.com/wp-content/uploads/2010/01/images_abc1280x960.png

2 Generating Data

The class as a whole will produce the images required for this assignment. Each student will collect, clean and submit **three** images each of American Sign Language gestures for the letters A - I, and K, for a total of **30** images.

2.1 Steps Involved in Data Collection

1. Familiarize yourself with American Sign Language gestures shown in Figure 1 for the letters A - I, and K (10 letters).
2. Ask a classmate to take **three** pictures at slightly different orientations for each letter gesture using your mobile phone. Keep in mind the following:
 - Ensure adequate lighting while you are capturing the images.
 - Use a white wall as your background.
 - Use your **right hand** to create the gestures.
 - Keep your right hand apart from your body and any other obstructions.
 - Avoid shadows.
3. Transfer the images to your laptop for cleaning.

2.2 Cleaning Data

To simplify the machine learning the task, you will convert the images into a standard format, as follows: Every image must have the same size and resolution: 56 x 56 pixels RGB. The hand must be in the center of the picture, which you can do by *cropping* the image appropriately. It is important to first crop your image to a *square* image (which will produce a large NxN pixel image). *After* the square crop, you must *resize* the image to 56x56 pixels. (If you don't crop to a square image, and then force a resize to a square 56x56 image, the image will become distorted, making the recognition problem harder.) You can use an online tools such as <http://picresize.com>, or native applications depending on your operating system, as below. For the resizing part, you can also use the Python Imaging Library (PIL) package in python.

Important: take a few pictures and then make sure that the cropping and resizing works properly, in case there is a problem with how you took the picture.

Mac OS

Use the **preview** app provided by MacOS. Holding down CMD + Shift will keep a square aspect ratio while selecting the hand area. Crop by selecting a region with the cursor, and selecting **Crop** from the tools menu. Then, select **Resize** from the tools menu to resize the image to 56x56 pixels.

Windows 10

Use the **Photos** app to edit and crop the image and keep the aspect ratio a square. Use the **Paint** to resize the image to the final image size of 56x56 pixels.

Linux

Use the **GIMP** app, **imagemagick**, or other tools of your choosing.

Note that the above steps are illustrative only. You need not follow these steps but following these will ensure that you produce a good quality dataset. Your pictures will be judged based on the quality of the images alone.

Please *do not* edit your photos in any other way. You should not need to change the aspect ratio of your image. You also should not digitally remove the background or shadows—instead, take photos with a white background and minimal shadows.

2.3 Accepted Images

Images will be accepted (for use by the whole class in Part 2 of this assignment) and graded based on the criteria below, with some illustration in Figure 2:

1. The final image should be size 56x56 pixels (RGB).
2. The file format should be a .jpg file and be named as described below.
3. The hand should be approximately centered on the frame.
4. The hand should not be obscured or cut off.
5. The photos follows the ASL gestures posted earlier.
6. The photos were not edited in any other way (e.g. no electronic removal of shadows or background).

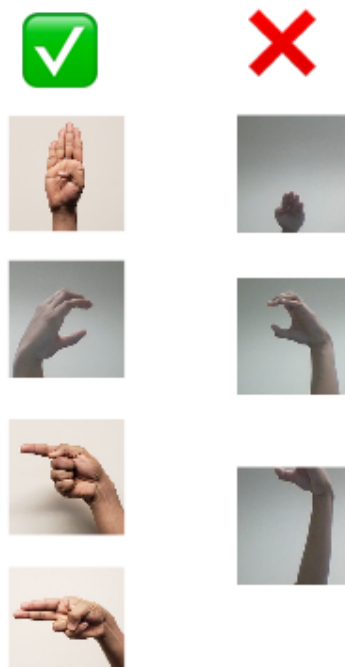


Figure 2: Acceptable Images (left) and Unacceptable Images (right)

2.4 Submission

Submit a zip file containing your 30 images, as follows:

1. Individual image file names must follow the convention of `student-number_Letter_number.jpg` (e.g. `100343434_A_1.jpg` for the first ‘A’ and `100343434_A_2.jpg` for the second, etc.).
2. Zip all the images together and name it with the following convention: `last-name_student-number.zip` (e.g. `Rose_100343434.zip`).
3. Submit the zipped folder.

Note that we will anonymize and combine the images that everyone submits. We will announce when the combined data set will be available for download.