

Sign Language Recognition

Team members

Adahan Yalçinkaya	21502369
Emre Sülün	21502214
Eray Şahin	21502758
Fuad Aghazada	21503691
Kazım Ayberk Tecimer	21502531

Description of the Data

The dataset that we are planning to use contains a set images of hand positions of 9 different people for American Sign Language (ASL) signs with different environment and lighting effects. For each alphabetical letter, the dataset contains around 200 images for each of 9 people (i.e. ≈ 1800 images per letter). More precisely, for each person distribution of the images is in the following table.

TABLE 1. Distribution of images per person

	Person # 1	Person # 2	Person # 3	Person # 4	Person # 5	Person # 6	Person # 7	Person # 8	Person # 9
# of images	3592	11431	7749	7101	5958	9784	10503	11526	5032

We are planning to use Cross Validation (CV) method, more specifically k-fold CV since our dataset is not so small. The value for k can be chosen as 5 or 10, but since there is no formal rule about configuring the k value, it can be modified during the implementation. We will normalize our image data by using various image processing techniques like contrast stretching, static and adaptive thresholding, edge detection and sharpening.

The dataset can be found on,

<http://empslocal.ex.ac.uk/people/staff/np331/index.php?section=FingerSpellingDataset>

Question That We Are Trying To Answer

ASL is a sign language that matches alphabetical letters to particular hand gestures. According to our research, only a *small* portion (1/312) of the US population uses ASL. To promote its use, in this project, we will investigate ways of interpreting ASL with the aid of machine learning and image processing principles. More specifically, given a live video stream (through the camera of the computer), by examining the position of hands, we aim to identify which letters are described according to ASL fingerspelling set. Each recognized letter is to be projected onto a window so as to inform the user simultaneously. In this way, the user will be able to write words onto the screen in an interactive manner.

In terms of the technical aspects, we are considering to use to Convolutional Neural Network (CNN), Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA).

Milestones

Until the progress report, we aim to model our dataset on at least one technique described in the question part. As a proof of concept, initially we will use few letters, 5 of the 26 letters in our dataset.