**Chapter 3**

**Data set**

**3.1 American Sign Language:**

American sign language is using to communicate between deaf community and normal community. However, there are only 2.5 million ~ 5.0 million speak which significantly limit the number of that they can easily communicate with [1].



**American Sign language Manual Alphabet [2].**

American Sign Language is implemented from French sign language which was introduced by Thomas Hopins Gallaudet in United States. ASL is similar to French sign language; Individuals who speak American Sign Language are able to effectively communicate in French Sign Language. A variation of American Sign Language also exits. Similarly, to English which is international language, but it has unique variations between English spoken in England, United States or Australian, there are separate difference that have changed in sign language (Stokoe, 2005).

**3.2 Characteristics of American Sign Language:**

* American Sign language is an entire visual-gestural dialect with its very own language structure, vocabulary, and linguistic structure.
* Like other sign language, it utilizes the hands, the body, and face looks (counting mouth developments) to express significance and the eyes to see meaning.
* Hand - to-hand connection is especially critical in ASL since it has no composed frame. There are, in any case, documentation frameworks that are utilized for recording signs on paper.
* ASL is separate from English and is unique from other signed languages. An example of the distinctiveness of signed languages from each other and from the surrounding spoken language(s) is that, although English is the shared spoken language of the U.S., Canada, and Britain, signers of ASL do not understand signers of British Sign Language (BSL).

**3.3. Statistics** **about sign language use in Canada:**

In Canada, Statistics Canada reports that as indicated by the 2006 Census 8,995 people revealed a gesture-based communication just like their primary language or one of their first languages, as gave beneath.

|  |  |
| --- | --- |
| American Sign Language | 2,485 |
| Quebec Sign Language | 730 |
| Sign languages, not included elsewhere | 5,780 |

**Table 1 : Statics about Sign Language as a Mother Tongue [3].**

In addition, Statistics Canada reports that as per the 2006 Census 43,090 people reported knowledge of a gesture-based communication, as provided below.

|  |  |
| --- | --- |
| American Sign Language | 11,110 |
| Quebec Sign Language | 730 |
| Sign languages, not included elsewhere | 5,780 |

**Table 2: Statics about Knowledge of Sign Languages[3].**

**3.4 Dataset and variables:**

I have captured images for my own dataset. This dataset was a collection of 36 which contain A to Z alphabet and 0 to 9 numbers digit. In my dataset consist of A to Z alphabet and 0 to 9 numbers where I have used right hand to capture 1200 images for specific alphabet and numbers. After that I implement code which convert flip image to right to left hand image. The height and width ratios vary significantly but average approximately 50X50 pixel. The A dataset contains over 100,000 images in gray scale color. Additionally, People who want to add their images to this dataset than they can add. Below figure shows a sample image of alphabet A, B,C respectively.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | |  |
| **A** | **B** | **C** | |

Figure 1: Sample images of my Data set.

|  |  |
| --- | --- |
| **Property** | **Description** |
| Alphabets and numbers | A to Z and 0 to 9 |
| Color | Gray Scale |
| Dimensions | 50x50 |
| Height | 50 pixels |
| Width | 50 pixels |
| File type | JPEG |

**Table 3: Dataset Description and Image property Histogram:**

The main principle goal of the histogram is to remove unnecessary background and noises, capture only the Region of Interest (ROI), which are the only useful figures in the image. **This is achieved via Skin Masking defining the threshold on RGB schema and then converting RGB colour space to grey scale image. Finally Canny Edge technique is employed to identify and detect the presence of sharp discontinuities in an image, thereby detecting the edges of the figure in focus.**

Histogram use to capture and store dataset images. Histogram stored in binary format in which I used pickle library to store histogram. Pickle is used for object serialization and deserialization. Pickling is process whereby python object hierarchy is converted into byte stream and unpickling is inverse operation. Once histogram is created after that I have started create my own dataset for my work which consist A to Z alphabets and 0 to 9 numbers image in gray scal.

1] Mitchell, Ross; Young, Travas; Bachleda, Bellamie; Karchmer, Michael (2006). "How Many People Use ASL in the United States?: Why Estimates Need Updating" (PDF). Sign Language Studies (Gallaudet University Press.) 6 (3). ISSN 0302-1475. Retrieved November 27, 2012.

2] Lifeprint.com. American Sign Language (ASL) Manual Alphabet (fingerspelling) 2007.

3] <http://www.manitoba.ca/index.html>