

Signing dataset for the Sinhala Sign Language

K. L. P. Prasad Liyanaarachchi¹, Dinithi Shakya¹, Thilina Herath¹, Nuwan Vithanage¹,
L.S.K. Udugama^{1,2,*}

¹Faculty of Computing, Sri Lanka Technological Campus, Padukka, Sri Lanka.

²The Open University of Sri Lanka, Nawala, Nugegoda, Sri Lanka

pushpikal@sltc.ac.lk, dinithis@sltc.ac.lk, thilinah@sltc.ac.lk, nuwanv@sltc.ac.lk,

*udugama@sltc.ac.lk, udugama@ou.ac.lk

ABSTRACT

Different Sign Languages are used by the hearing impaired people (HIP) to express their ideas, thoughts, and feelings among the people in society. Sinhala Sign Language (SSL) is one of the Sign Languages used by the HIP in Sri Lanka. The main problem faced by the HIP is the communication difficulty with the people who do not understand the sign language. One solution to this problem is developing an automated sign language interpreter. For this purpose, it is necessary to have a proper signing dataset. Although some research works have been published on constructing datasets for various Sign Languages around the world there are no signing datasets in video format for SSL. Therefore, this research is to construct a benchmark signing dataset for the SSL.

In this research, only the published signs of SSL will be used to build the signing dataset and it will be limited to the classroom domain. Further, it considers the vocabulary of sign languages from a linguistic point of view. There are three components in the sign language such as isolated signs (Static signs), continuous signs and annotations. Based on previous research, this dataset will be consisting of three parts in folder structure such as Video Bank, Frames and Annotations. To collect signing data, it needs a well-organized lab setup. The lab has four video cameras and Kinect sensors to capture signing. To train systems such as CNN and ANN, these cameras need to capture accurate video data from different angles. Therefore, we need more than one camera as all movements of a sign should be captured without losing any information. As the initial step, this lab is set up to do the experiments to find a better orientation of the cameras and they will be synchronized when capturing videos. Further, the dataset will be fine-tuned according to various factors: performance, size and accuracy of recognizing the sign.

The purpose of this research is to construct a signing dataset for SSL which will be beneficial to those who wish to build systems for the HIP for communicating among the people in society. Further, this dataset will be available to anyone who conducts research in SSL.

Keywords: Sign Language, Lab setup for signing, signing dataset, Sinhala Sign Language.