MAIS 202 Final Project Deliverable 1: Sign Language Translator

Gohar Saqib Fazal

I. Dataset

I was first inspired to make the Indo-Pakistani sign language translator when I read a research paper on how computer vision analysis can revolutionize the lives of the deaf population [1]. Pakistan has a deaf population of 0.24 million, which is approximately 7.4% of the overall disabled population in the country and this language is used by at least several hundred thousand deaf signers (2003) [2]. Unlike American sign language this language is in its rudimentary stages. In order to make it more recognized in the world I will be using alphabetical data in Urdu language which is easily available online and then building the dataset myself. I will be using the concept of supervised learning to input the a group of labelled data (photos of me signing the data) into the computer. The computer will then store the labelled data in a database and find similarities about objects of the same label in a live video feed transmission.

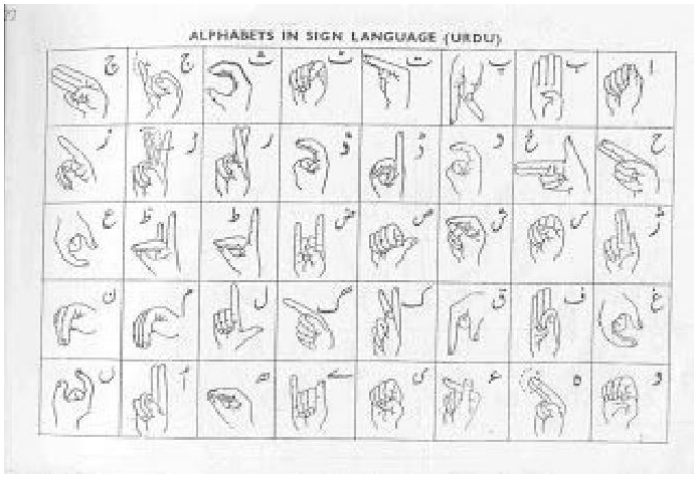

Fig 1. A picture of the alphabets and their corresponding hand gestures used in Urdu.

II. Methodology

a. Data Preprocessing

Since the data involves direct photographs of me there is not much data processing involved. During the process, it will be ensured that the picture maintains a specific quality by using a focused 12 Megapixel camera. Unfocussed pictures not meeting the requirements will be discarded and new pictures taken. To increase accuracy I will be taking ten photos for each letter and then putting every letter in its own directory.

b. Machine Learning Model

In order to make the translator I will be training my own convoluted neural network and using computer vision. I will be using Tensor Flow or Pytorch for the purpose of the application (whichever is more feasible). The predicted model for the project will be required to at least have an 80% accuracy. The translator can also be implemented using K-nearest neighbors classification (KNN) but with lesser accuracy if the convoluted neural network fails to work due to some reason.

c. Final Conceptualization

I aim to present the final model for the translator as either a web application or a mobile application (even though I am not that experienced). Once the model is trained I will be needing a back-end and a front-end service in order to make an application. I will be wrapping my model in a Flask app (Python) for the backend and JavaScript for the frontend. This will allow for the incorporating of real time data from the users' webcam every 200ms and then displaying the results [3].

References:

[1]B. Hassan, M. Farooq, A. Abid and N. Sabir, "Pakistan Sign Language: Computer Vision Analysis & Recommendations", VFAST Transactions on Software Engineering, vol. 9, no. 1, p. 1, 2015. Available: 10.21015/vtse.v9i1.386 [Accessed 1 October 2019].

[2]"Wayback Machine", Web.archive.org, 2019. [Online]. Available: https://web.archive.org/web/20120315001919/http://www.amic.org.sg/ict/external/awards/0202a2_l59attachment2.pdf. [Accessed: 01- Oct- 2019].

[3]"How we used AI to translate sign language in real time", Medium, 2019. [Online]. Available: https://medium.com/@coviu/how-we-used-ai-to-translate-sign-language-in-real-time-782238ed6bf. [Accessed: 01- Oct- 2019].