Real-Time Multilingual Sign Language Translation into Text

Birundha Jayaraman – birundha2.jayaraman@live.uwe.ac.uk
Lokesh Krishnasamy – lokesh2.krishnasamy@live.uwe.ac.uk
Rasagna Bellam – rasagna2.bellam@live.uwe.ac.uk

Keywords: Multilingual Sign Language, Deep Learning, Computer Vision, Neural Architecture Search Network, Long Short-Term Memory, Sequence-to-Sequence Model

Description:

Scope:

The scope of this research is to develop a multilingual sign language translation model combining deep learning and computer vision techniques to improve accessibility for persons with hearing loss. The model developed will have the potential to be integrated into real-time sign language translation applications. We will also look at the ethical and legal challenges of using such a model.

Problem Statement:

The communication barrier between people with hearing impairment and non-sign language fluent people makes efficient contact difficult. To translate sign language accurately and rapidly into text in various languages, a real-time multilingual sign language translation model is required. A model like this would improve social inclusion and accessibility for sign language users by allowing them to communicate across various sign languages used in different countries.

Rationale:

Prior research has looked at sign language recognition and translation models [7], but most of them have limits in terms of accuracy, speed, lack of results and multilingualism [6]. Moreover, such models frequently need manual interpretation, which may be time-consuming and error-prone. The construction of a real-time multilingual sign language translation model is now possible because of recent advances in deep learning and computer vision. The model that we will develop will use NASNet [2], LSTM [3], and the sequence-to-sequence model [4], which have demonstrated promising results in image recognition and natural language processing.

Aim & Objectives:

The research aims to develop a real-time multilingual sign language translation model that identifies and interprets sign language gestures into text in many languages. The objectives of the research are:

- To do a literature review on sign language recognition and translation.
- In the Computer Vision stage of the model: boost performance by using NASNet pre-trained on ImageNet.
- In the NLP stage of the model: increase prediction accuracy using LSTM.
- We will use the sequence-to-sequence paradigm to handle tasks like machine translation and text summarization effectively.
- Assessing the model performance in terms of accuracy and efficiency.
- Identifying ethical and legal considerations associated with the model use.
- Investigating potential use cases for the model in various contexts.

Methodology:

Proposed Approach:

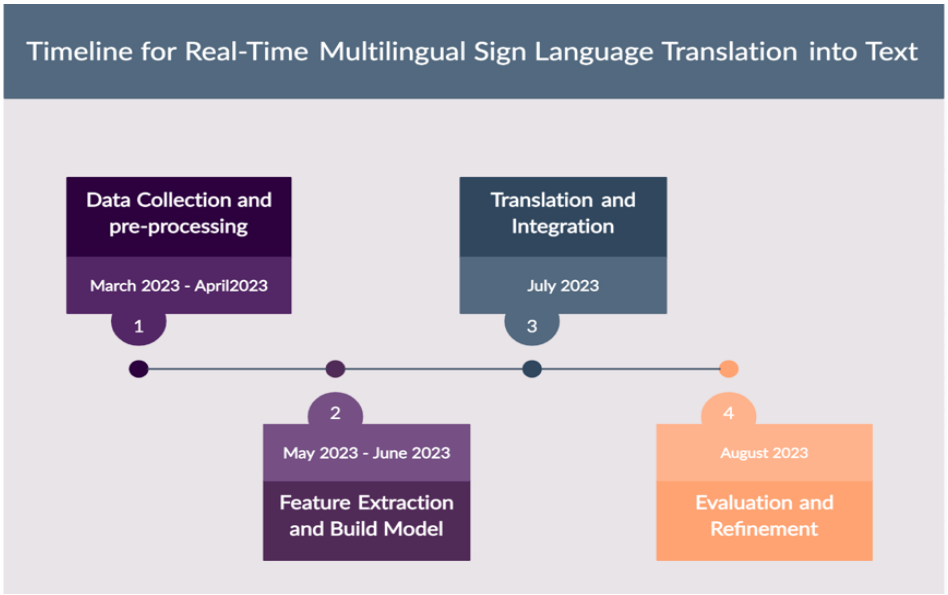
Our proposed approach to building a real-time multilingual sign language translation model using deep learning and transfer learning [5] involves several steps. The first step is to pre-process the videos obtained from the SP-10 dataset [1] using techniques such as optical character recognition, object

detection, and image segmentation. The second step is feature extraction, which involves using a pre-trained NASNet model [2] to recognize objects in the videos. The third step is sequence modelling using LSTM [3] to generate a sequence of feature vectors that model the temporal dependencies in the videos. Finally, a sequence-to-sequence model [4] is used to translate the sign language into text. A parallel corpus of sign language videos and their associated translations is required for the training of a sequence-to-sequence model for sign language translation, which offers the essential data for the model to learn from. Teacher forcing and beam search decoding approaches are used throughout the training phase to increase the accuracy of the model's output by directing it towards the proper translation during the decoding step.

Project Management Method:

We are going to use Agile Methodology to develop our project. Following the Agile manifesto, to enhance flexibility and collaboration in the group, we have already started to use Zotero, Jira and Bitbucket for our project.

Project Timeline:



References:

1. Yin, A., Zhao, Z., Jin, W., Zhang, M., Zeng, X., & He, X. (2022). MLSLT: Towards Multilingual Sign Language Translation. In 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 5099-5109). New Orleans, LA, USA: IEEE. doi: 10.1109/CVPR52688.2022.00505.

2. Adedoja, A., Owolawi, P.A., & Mapayi, T. (2019) Deep Learning Based on NASNet for Plant Disease Recognition Using Leave Images. In Proceedings of the International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD), 1-5. doi: 10.1109/ICABCD.2019.8851029.

3. Liu, T., Zhou, W., & Li, H. (2016). Sign language recognition with long short-term memory. In 2016 IEEE International Conference on Image Processing, ICIP 2016, Phoenix, AZ, USA, September 25-28, 2016, pages 2871–2875. IEEE. doi: 10.1109/ICIP.2016.7532884.

4. Arvanitis, N., Constantinopoulos, C., & Kosmopoulos, D. (2019). Translation of Sign Language Glosses to Text Using Sequence-to-Sequence Attention Models. 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), 296-302. doi: 10.1109/SITIS.2019.00056.

5. Thakar, S., Shah, S., Shah, B. and Nimkar, A.V. (2022) 'Sign Language to Text Conversion in Real Time using Transfer Learning', in 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT), Bangalore, India, pp. 1-5. doi: 10.1109/GCAT55367.2022.9971953.

6. Huynh, T.-L., Nguyen-Ngoc, K.-N., Chu, C.-B., Tran, M.-T., & Le, T.-N. (2022). Multilingual Communication System with Deaf Individuals Utilizing Natural and Visual Languages, 683-688. doi: 10.1109/RIVF55975.2022.10013851.

7. Singh, J. and Singh, D. (2022) 'A Comprehensive Review on Sign Language Recognition Using Machine Learning', in Proceedings of the 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), pp. 1-6. doi: 10.1109/ICRITO56286.2022.9965118.