Sign Language Translation for ISL
Using Procedural Learning and Optimal Transport

Aritra Ambudh Dutta (230191)
Ashish Upadhyay (230231)
Siddhant Shekhar (231003)

under the guidance of
Prof. Ashutosh Modi

June 20, 2025

Abstract

This project explores the use of procedural learning and optimal transport for improving Indian Sign Language (ISL) detection at the word level. The focus of this work has been on understanding the concepts from the OPEL (Optimal Transport Guided ProcedurE Learning) framework, which aligns video frames by treating them as probability distributions and using optimal transport to learn the step-wise procedure present in video sequences. Through this, we aim to understand how such alignment-based learning can be applied to sign language data, where the temporal structure is crucial.

To begin with, I studied the mathematical foundations of optimal transport and how it was used in the OPEL paper. I explored how the method avoids hard frame-to-frame alignment and instead learns soft, distribution-based alignment, which can be more robust for sign language understanding. Several resources, including video explanations and research papers, were used to deepen my understanding of these methods.

In parallel, we curated and cleaned a large-scale dataset of Indian Sign Language videos. The dataset contains approximately 1.5 lakh videos, each representing a different collection of words in ISL. These videos were collected and filtered to ensure quality, diversity, and relevance to the task. This dataset will be used in upcoming model development and training phases.

I am exploring the SHuBERT framework, a self-supervised model for sign language, incorporating its multi-stream visual processing and masked cluster prediction techniques alongside OPEL's procedural alignment methods. Future work will integrate insights from both models, training first on standard sign language datasets and then adapting to Indian Sign Language using our curated video data.

Special thanks to my mentor, Sanjeet, for his continuous guidance and support throughout this project.

References

- **OPEL: Optimal Transport Guided ProcedurE Learning**
  https://proceedings.neurips.cc/paper_files/paper/2024/hash/6e4b14e76d0d4f42a9dff031a7a8417b-Abstract-Conference.html

- **SHuBERT: Self-Supervised Sign Language Representation Learning via Multi-Stream Cluster Prediction**
  https://arxiv.org/abs/2411.16765