

Work Report

The initial task for text-to-sign language conversion involves creating an English-to-HamNoSys dataset. Using web scraping techniques, around 57,000 English sentences from various online sources were collected, in addition to certain publicly available datasets. Since manual conversion takes time, an automated system was created to speed up the procedure. This automation system was built using an existing English word-to-HamNoSys dictionary.

Through automation, a parallel corpus comprising 59,000 English sentences and their corresponding HamNoSys translations was generated. Tokenizers for HamNoSys and English were subsequently created. After that, a transformer architecture was used to build a neural machine translation model, which was then trained on the parallel corpus. After training for 100 epochs, the model's loss was 1.09.

But because there wasn't enough data, the model didn't perform well during testing.

Traditionally, text-to-sign translation involves converting text to gloss, gloss to HamNoSys, HamNoSys to XML code (SiGML), and SiGML to animation. To streamline this process, we implemented a system to convert HamNoSys to SiGML. For converting SiGML to animation, we plan to integrate our system with the JA Signing App developed by the University of East Anglia.

Additionally, we explored using animation software like Blender for text-to-sign translation. We have now completed creating digit sign animations using Blender.