

Including Signed Language in Natural Language Processing

Kayo Yin, Amit Moryossef, Julie Hochgesang, Yoav Goldberg, Malihe Alikhane

Authors in ACL'21 in their position paper wants the NLP community to expand and include Sign Language, a primary means of communication for many deaf and hard hearing people, in Natural Language Processing as it not only follows all the fundamental linguistic properties of natural language but also have a high social and scientific impact. In the paper, the authors first showed that signed language exhibits all fundamental linguistic properties, reviewed the limitations of current Signed Language Processing Techniques, and finally advocated towards adopting an efficient tokenization method by building signed language databases.

Authors posited that sign language like other natural languages also consists of phonological (hand configuration, palm orientation/placement, eye and head movements), morphological, syntactic and semantic structure. Authors also pointed out that though sign-language takes long to produce, the information transmitted at a given time is the same as sign-language as it simultaneously makes use of multiple visual cues to convey series of information, or facial cues to modify parts of speech thereby reducing time. People also take help of referencing, fingerspelling (spelling letters to form words), poses, role shifts(by shifting places) to convey their thoughts.

Until now, research in signed language was focused mostly on videos (because of their high dimensionality it is very difficult to process), poses or glossing (interpreting sign language per sign). Therefore the steps taken were detection (whether video has sign language present), extracting sign language, segmentation, recognition and then translation. Sign language datasets that are created till now are also based on videos or images. Researchers also exploited the distribution of phonemes to identify sign language. But time has come to expand natural language techniques and for that authors suggest building a proper pipeline consisting of tokenization (defining lexical units), syntactic analysis (understanding morphological features), named entity recognition(through fingerspelling), and coreference resolution (very important to find relations).

I personally find the paper very inspiring, the topic has the potential to bridge the gap in interpreting sign language by everyone in real time and see the world through a whole different modality. Sign Language also has

the potential to understand historical documents. It was very interesting to see so many similarities between the two languages.

But I also feel, as there is no standard sign language, and different modality, we will face various challenges but hope to see numerous progress in including sign language in nlp and cv.

To conclude, as sign-language is very similar to other natural languages, we should not ignore it and rather include it for the betterment of society.