

Using an LLM to Turn Sign Spottings into Spoken Language Sentences

Ozge Mercanoglu Sincan, Necati Cihan Camgoz, Richard Bowden

Sign Language Translation (SLT) is a challenging task that aims to generate spoken language sentences from sign language videos. In this paper, we introduce a hybrid SLT approach, Spotter+GPT, that utilizes a sign spotter and a powerful Large Language Model (LLM) to improve SLT performance. Spotter+GPT breaks down the SLT task into two stages. The videos are first processed by the Spotter, which is trained on a linguistic sign language dataset, to identify individual signs. These spotted signs are then passed to an LLM, which transforms them into coherent and contextually appropriate spoken language sentences. The source code of the Spotter is available [here](#) 1.