

Paper Title:

Sign Language Recognition and Translation: A Multi-Modal Approach using Computer Vision and Natural Language Processing

Paper Link:

<https://aclanthology.org/2023.ranlp-1.71/>

1 Summary**1.1 Motivation**

The paper is motivated by the significant global population facing hearing loss and the limited accessibility of sign language, especially in the context of nearly 2.5 billion individuals estimated to be affected by hearing impairment by 2050. The authors emphasize the need for effective non-verbal communication methods and identify the challenges faced by those with hearing loss.

1.2 Contribution

The primary contribution of the paper lies in proposing a multi-modal approach, Sign-to-Text (S2T), for sign language recognition and translation. The authors introduce a combination of Computer Vision and Natural Language Processing (NLP) techniques to address the limitations of existing methods. The integration of NLP, specifically Autocorrect, Context Awareness, and Machine Translation, serves as a novel contribution to enhance the accuracy and robustness of sign language recognition.

1.3 Methodology

The methodology involves two iterations of S2T. The initial version employs Computer Vision for gesture classification but faces limitations in accurately classifying all ASL signs. The shortcomings of Computer Vision are addressed in the second iteration by incorporating NLP techniques. Autocorrect algorithms, including Needleman-Wunsch, TextBlob, Minimum Edit Distance, and Damerau-Levenshtein, are introduced to improve gesture classification accuracy. Additionally, the paper explores Machine Translation using the Transformer model and IBM Models to enhance the translation of sign language into text.

1.4 Conclusion

In conclusion, the paper highlights the success of the multi-modal approach in improving sign language recognition and translation. Autocorrect significantly enhances error correction in gesture classification, setting the stage for more accurate translation. The Transformer model demonstrates promising results in translation quality, but challenges remain in the evaluation of Statistical Machine Translation (SMT) methods. The authors propose future work, including refining Autocorrect, exploring advanced context awareness methods, and integrating BERT models for improved predictions and hybrid approaches. The paper underscores the potential for advancements in sign language accessibility through innovative technologies.

2 Limitations

2.1 First Limitation

The paper acknowledges subpar performance in machine translation, specifically with the Transformer model scoring lower than expected on the BLEU metric. However, the limitations lie in the limited evaluation scope. The chosen dataset for evaluation might not fully represent the complexities of translation tasks. A broader evaluation on diverse datasets or benchmark corpora, such as those from past WMT conferences or OPUS, would provide a more comprehensive understanding of the MT capabilities of the proposed system.

2.2 Second Limitation

While the paper introduces context awareness as a valuable component, the performance of the implemented SpaCy-ContextualSpellCheck pipeline is noted as not substantial enough for reliable use. The limitations of this approach suggest room for improvement in the methodology related to context awareness. Exploring alternative algorithms or incorporating more advanced techniques, such as the Viterbi algorithm or direct implementation of BERT models, could enhance the overall performance of context awareness in autocorrect and contribute to better accuracy in sign language recognition.

3 Synthesis

The paper presents Sign-to-Text (S2T), a novel approach to enhance American Sign Language (ASL) recognition and translation by combining Computer Vision and Natural Language Processing (NLP). The multi-modal system addresses the limitations of pure Computer Vision by introducing autocorrect, context awareness, and machine translation in conjunction with gesture classification. The proposed methodology significantly improves the accuracy of hand gesture recognition, leading to more reliable sign language translation. While demonstrating promising results, the paper also acknowledges areas for future refinement, emphasizing the potential of this multi-modal approach to advance communication accessibility for individuals with hearing impairments.