Gloss Attention for Gloss-free Sign Language Translation

1. Limitations and Potential Negative Effects

Our work does not have apparent negative effects, but since facial expressions are also an important part of sign language, it can convey a wealth of information, and the existing methods based on RGB video sequences may infringe on users' privacy. In order to protect the privacy of users, in the future, we can try to develop some facial blur methods that do not affect translation performance, and sign language translation methods based on skeletal sequences. The main limitation of our method is that it still lags in translation performance relative to the state-of-the-art gloss-supervised sign language translation methods. In the future, we can try to explore whether more data without gloss can be used to achieve the same translation performance as data with gloss, because it is much simpler to obtain sign language translation data without gloss annotation. Another direction worth exploring is whether semi-supervised learning using a small amount of gloss-annotated sign language translation data can achieve the same or close translation performance as the full gloss data.

2. Qualitative Comparison.

Although the qualitative analysis table in the main text is a vector and can be enlarged for viewing, we have included a larger version in the appendix for easier viewing. We present 3 example translation results generated by our GASLT model and TSPNet model in Table 1 for qualitative analysis. In the first example, our model produces a very accurate translation result, while TSPNet gets the date wrong. In the second example, our model ensures that the semantics of the sentence has not changed by using the synonym of "warnungen" (warnings) such as "unwetterwarnungen" (severe weather warnings), while TSPNet has a translation error and cannot correctly express the meaning of the sign language video. In the last example, it can be seen that although our generated results differ in word order from the ground truth, they express similar meanings. However, existing evaluation metrics can only make relatively mechanical comparisons, making it difficult to capture these differences. We provide the full translation results generated by our proposed model in the supplementary material.

References

[1] DONGXU LI, Chenchen Xu, Xin Yu, Kaihao Zhang, Benjamin Swift, Hanna Suominen, and Hongdong Li. TSPNet: Hierarchical Feature Learning via Temporal Semantic Pyramid for Sign Language Translation. In *Advances in Neural Information Processing Systems*, volume 33, pages 12034–12045. Curran Associates, Inc., 2020. 2

Table 1. Comparison of the example gloss-free translation results of GASLT and the previous state-of-the-art model. We highlight correctly translated 1-grams in blue, semantically correct translation in red.

Ground Truth:	und nun die wettervorhersage für morgen donnerstag den siebzehnten dezember .
	(and now the weather forecast for tomorrow thursday the seventeenth of december .)
TSPNet [1]:	und nun die wettervorhersage für morgen donnerstag den sechzehnten januar .
	(and now the weather forecast for tomorrow thursday the sixteenth of january .)
Ours:	und nun die wettervorhersage für morgen donnerstag den siebzehnten dezember .
	(and now the weather forecast for tomorrow thursday the seventeenth of december .)
Ground Truth:	es gelten entsprechende warnungen des deutschen wetterdienstes .
	(Appropriate warnings from the German Weather Service apply .)
TSPNet [1]:	am montag gibt es hier und da schauer in der südwesthälfte viel sonne .
	(on monday there will be showers here and there in the south-west half, lots of sun .)
Ours:	es gelten entsprechende unwetterwarnungen des deutschen wetterdienstes .
	(Appropriate severe weather warnings from the German Weather Service apply .)
Ground Truth:	morgen reichen die temperaturen von einem grad im vogtland bis neun grad am oberrhein .
	(tomorrow the temperatures will range from one degree in the vogtland to nine degrees on the upper rhine .)
TSPNet [1]:	heute nacht zehn grad am oberrhein und fünf grad am oberrhein .
	(tonight ten degrees on the upper rhine and five degrees on the upper rhine.)
Ours:	morgen temperaturen von null grad im vogtland bis neun grad am oberrhein .
	(tomorrow temperatures from zero degrees in the vogtland to nine degrees on the upper rhine .)