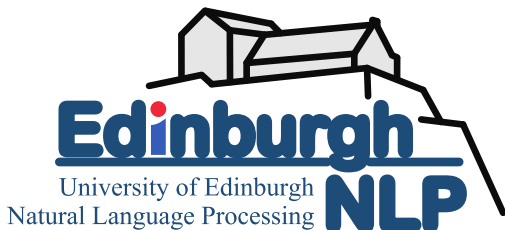


SLTUnet: A Simple Unified Model for Sign Language Translation

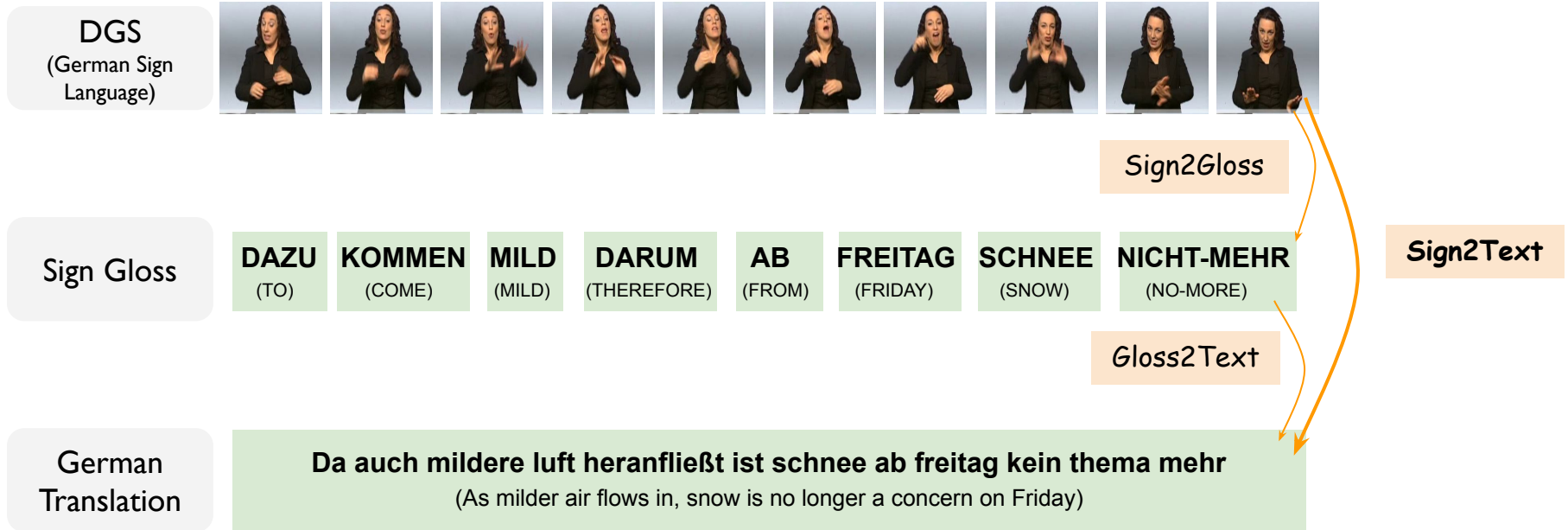
Biao Zhang¹, Mathias Müller², Rico Sennrich^{2,1}

¹University of Edinburgh ²University of Zurich



University of
Zurich^{UZH}

Sign Language Translation



Why Unified Modeling for SLT?

✘ SLT suffers

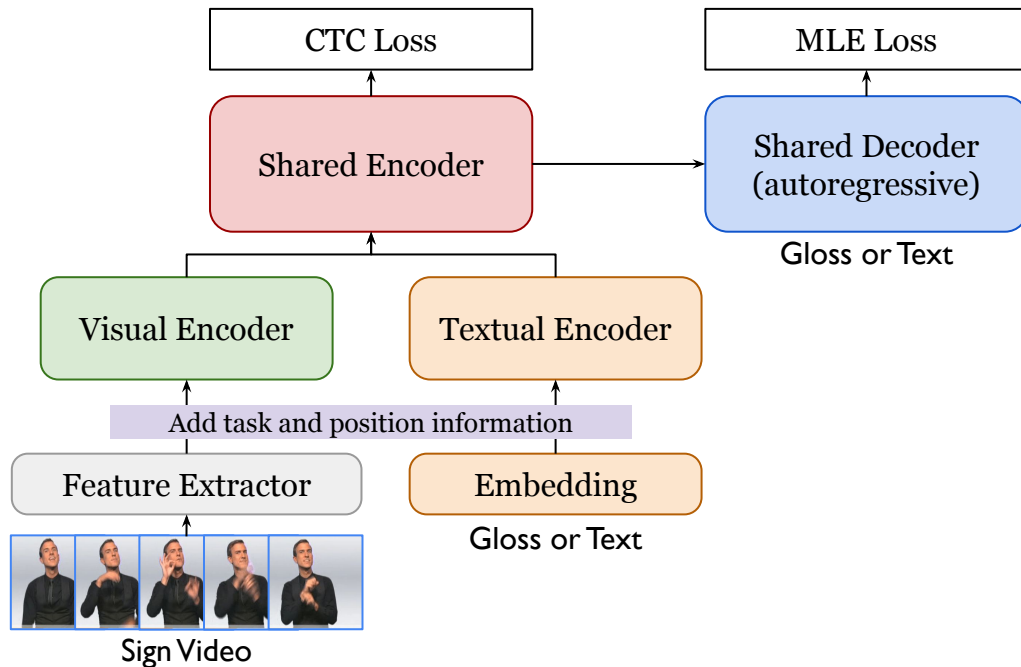
- Modality gap between video and text
- Lack of training data

✔ Unified modeling offers

- Implicit regularization with better data efficiency
- Cross-task & cross-modality transfer learning
- More universal representations
- The use of external sources

Jaegle et al., 2022; Zheng et al., 2021; Devlin et al., 2019; Zhang et al., 2020; Liang et al., 2022;

SLTUnet: The Proposed Unified Model



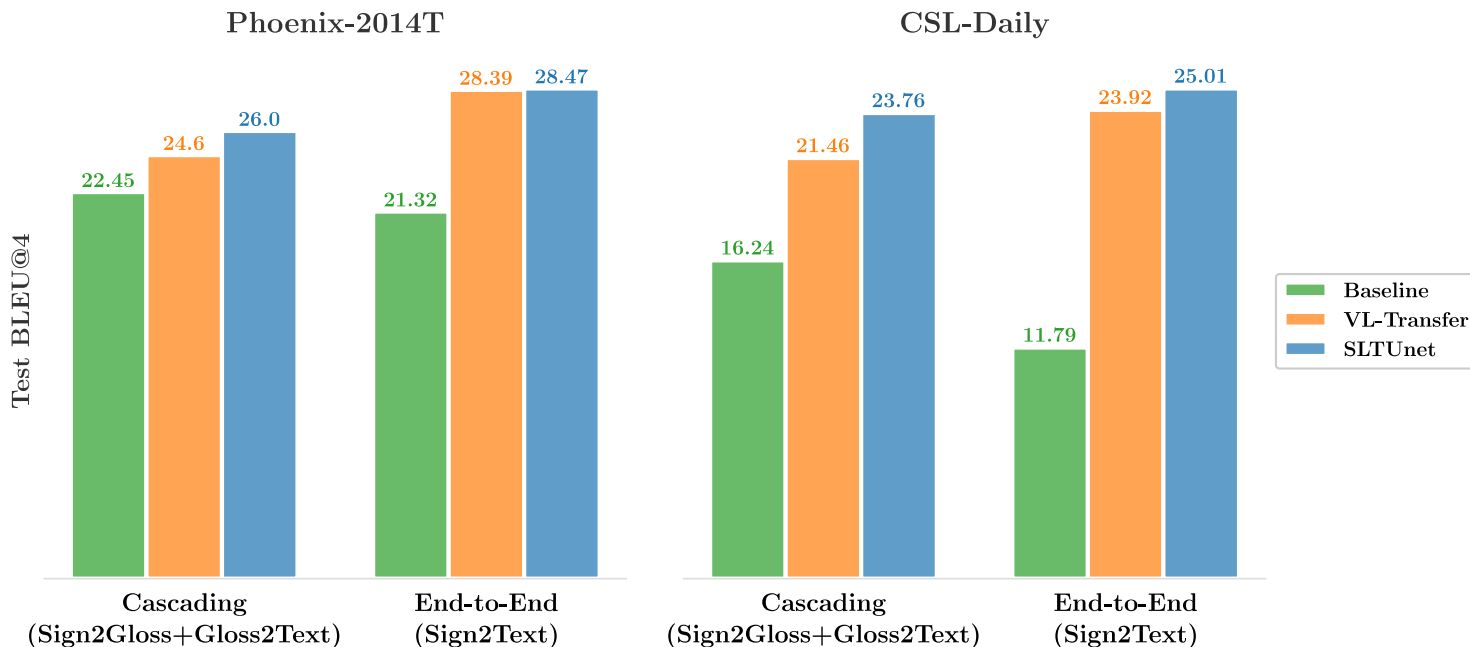
Sign2Text (video \rightarrow text)
 $\alpha \text{CTC}(\text{gloss}) + \text{MLE}(\text{text})$

Sign2Gloss (video \rightarrow gloss)
 $\text{MLE}(\text{gloss})$

Gloss2Text (gloss \rightarrow text)
 $\text{MLE}(\text{text})$

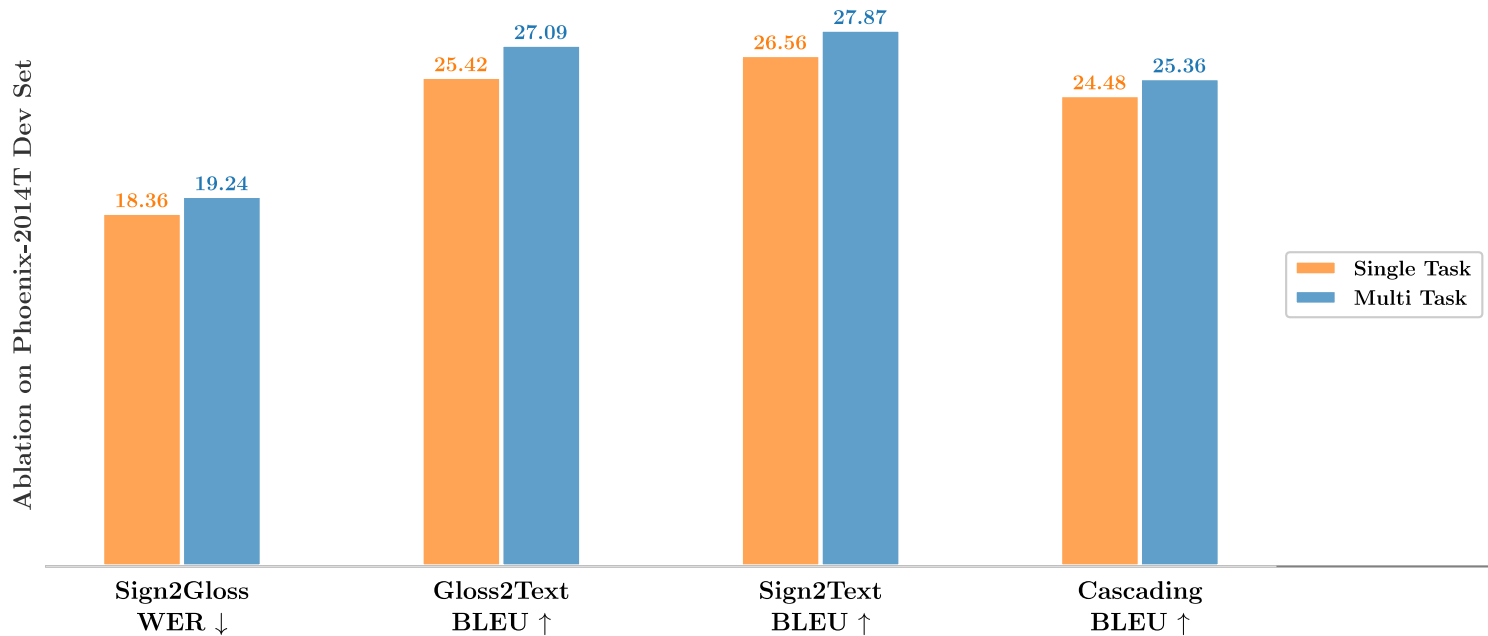
Machine Translation (source \rightarrow target)
 $\text{MLE}(\text{target})$

SLTUnet Achieves Promising Translation Performance



Phoenix-2014T: German sign language to German; CSL-Daily: Chinese sign language to Chinese
Baseline: (Camgoz et al., 2020) VL-Transfer: (Chen et al., 2022). MT data come from MuST-C

Multi-task Learning Benefits Translation

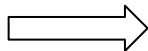


DGS3-T: A New SLT Testbed (Public DGS Corpus)

9 signers
1,085 glosses
Weather domain
7,096 #train
German Sign Language
video / gloss
German text

SLTUnet: **28.47** BLEU

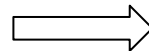
Phoenix-2014T
(Camgoz et al., 2018)



10 signers
2,000 glosses
Daily Life domain
18,401 #train
Chinese Sign Language
video / gloss
Chinese text

SLTUnet: **25.01** BLEU

CSL-Daily
(Zhou et al., 2021)



330 signers
8,580 glosses
Diverse domains
60,306 #train
German Sign Language
video / gloss
German text

SLTUnet: **2.82** BLEU

DGS3-T
(Ours, 2022)

Summary

- We propose **SLTUnet**
 - A unified model supporting multiple SLT-related tasks.
 - Greatly improves SLT, obtaining promising results on Phoenix-2014T and CSL-Daily.
- We propose to use a new SLT testbed, **DGS3-T**
 - Larger scale, covering more signers, richer topics/domains and more glosses.
- DGS3-T represents more practical challenges
 - SLTUnet performs poorly on DGS3-T
 - Previous progress might be overestimated

Paper: https://openreview.net/forum?id=EBS4C77p_5S

