ICASSP2021 TORONTO Canada June 6-11, 2021 Metro Toronto Convention Centre

Paper ID: 1981

Modeling Homophone Noise for Robust Neural Machine Translation

Wenjie Qin[1], Xiang Li[2], Yuhui Sun[2], Deyi Xiong[3], Jianwei Cui[2], and Bin Wang[2]

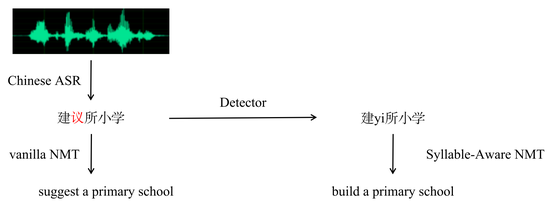
[1]School of Computer Science and Technology, Soochow University [2]Xiaomi AI Lab [3]College of Intelligence and Computing, Tianjin University

Abstract

We propose a robust neural machine translation (NMT) framework to deal with homophone errors. The framework consists of a homophone noise detector and a syllable-aware NMT model. The detector identifies potential homophone errors in a textual sentence and converts them into syllables to form a mixed sequence that is then fed into the syllable-aware NMT. Extensive experiments on Chinese→English translation demonstrate that the proposed method not only significantly outperforms baselines on noisy test sets with homophone noise, but also achieves substantial improvements over them on clean texts.

Motivation

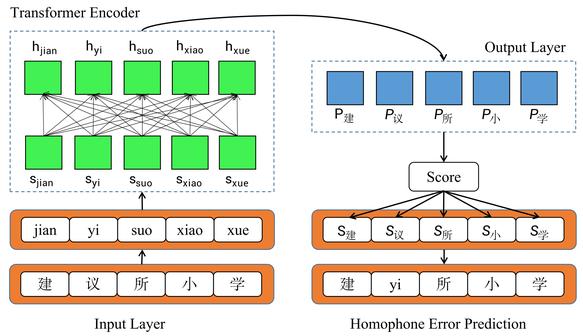
Despite remarkable progress made in NMT recently, most NMT systems are still prone to translation errors caused by noisy input sequences. One common type of input noise is homophone noise, where words or characters are mis-recognized as others with the same or similar pronunciation in ASR.



Detector

We predict the each character in each original textual sentence based on corresponding clean syllable sequence during training. For inference, we compute the log-likelihood score (LLS) of each output token based on the whole syllable sequence as follows:

$$LLS(x_i) = logP(x_i|S;\theta), \quad (1)$$



Effect of Noise Ratio

Model	0.1	0.2	0.3	0.4	0.5
Baseline	32.17	22.00	14.86	10.19	7.09
Li et al.	43.32	42.96	42.51	42.07	41.62
Liu et al.	42.34	41.97	41.92	41.52	41.38
Our SANMT	**44.58**	**44.30**	**44.34**	**43.05**	**42.34**

We reported the average BLEU scores of all systems on ANTs with a noise ratio varying from 0.1 to 0.5. Our method performs the best at all settings.

Robustness to Real-World Noise

Clean Utterance	请拼写它
Noisy Transcript	请拼写他
Li et al.	Please spell him
Liu et al.	Please spell him
Output of Detector	请 pin xie ta
Output of SANMT	Please spell it
Clean Utterance	听医生的建议
Noisy Transcript	听一生的建议
Li et al.	Listen to the doctor's advice
Liu et al.	Suggestions on life
Output of Detector	听 yi 生的建议
Output of SANMT	Listen to doctors' advice

The first example contains a third-person pronoun error, which is common in Chinese ASR as "他", "她", and "它" are all pronounced as "tā". The erroneous homophone word "一生" in second example is a noun which has a different meaning from the original homophone "医生".

Conclusions

We have presented a novel framework composed of a homophone error detector and an SANMT model to cope with homophone noise. Experimental results show that our method not only achieves substantial improvement over previous robust NMT both on the noisy test sets, but also outperforms the baselines on the clean test sets.