1 ASR Improvement

Our efforts to enhance the Automatic Speech Recognition (ASR) system have centered around extending the training duration and decreasing the token num for the output.

We have extended the Automatic Speech Recognition (ASR) training duration to 2 days, during which a total of 50 epochs were calculated.

1.1 Wav2vec WER Score before and after Improvement

Type	Beam Size	Wav2vec WER before	Wav2vec WER after
Test Set	5	149.43	9.38

2 Speech Translation

We successfully integrated the Automatic Speech Recognition (ASR) and Machine Translation (MT) systems for Cascaded Speech Translation (ST), resulting in the following overall performance.

Part	Beam size	Score Type	Score Result
ASR	5	WER	9.38
MT	4	BLEU	28.7
ST		BLEU	18.6

Now our current agenda involves directing our efforts towards Advanced Improvements to further enhance the Speech Translation score.