A Multi-Modal Chinese Poetry Generation Model

Dayiheng Liu, Quan Guo, Wubo Li and Jiancheng Lv
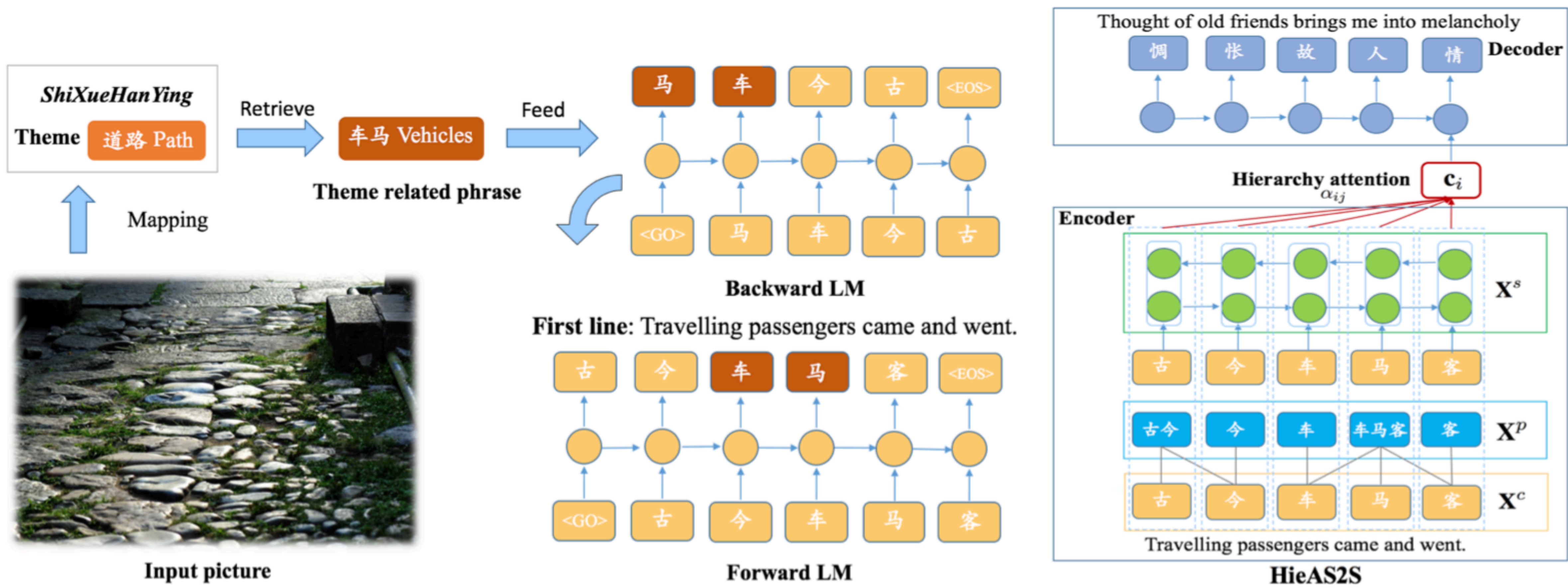
Machine Intelligence Laboratory

College of Computer Science, Sichuan University

Chengdu 610065, P.R. China

losinuris@gmail.com

Introduction & Model

This paper proposed a three-stage multi-modal Chinese poetry generation approach. Given a picture, the first line, the title and the other lines of the poem are successively generated in three stages. We propose a hierarchy-attention seq2seq model which can effectively capture character, phrase, and sentence information between contexts and improve the symmetry delivered in poems. In addition, the LDA model is utilized for title generation and improve the relevance of the whole poem and the title.

Experiments

Data

Chinese Poetry Corpus *ShiXueHanYing* KB *PoetryImage* Dataset

- 149,524 Poems
- 40,000+ thematic phrases
- 40,000+ Manually labeled Images

Results

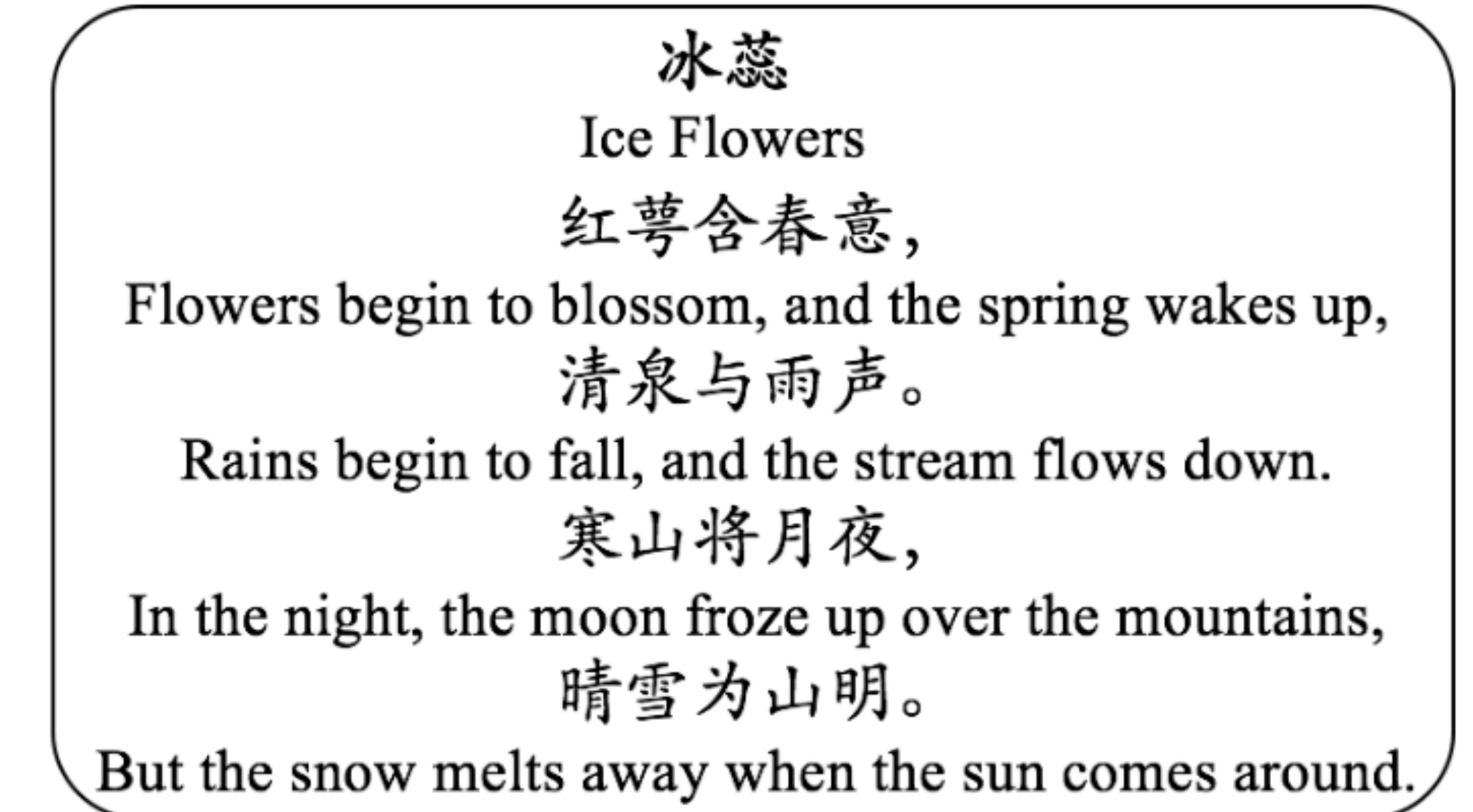| Approach	Metrics	
	BLEU-2	RHYTHM
baseline	26.726	0.824
AS2S	27.458	0.866
HieAS2S-tile	**29.991**	**0.892**
HieAS2S-concat	28.171	0.876
Positive-groundtruth	29.095	1.000
Negative-groundtruth	12.062	1.000

Automatic Evaluation corresponds to the upper rows; **Human Evaluation** corresponds to the table below.

Method	Poeticness	Fluency	Meaning	Coherence	Overall
AS2S	3.62	3.07	2.73	3.12	3.13
Ours	3.87	3.24	2.85	3.24	3.30
Human-written	4.07	3.43	3.58	3.71	3.69

Examples

(a) The outputs of *image recognition module* for the user-uploaded image.

冰蕊
Ice Flowers
红萼含春意，
Flowers begin to blossom, and the spring wakes up,
清泉与雨声。
Rains begin to fall, and the stream flows down.
寒山将月夜，
In the night, the moon froze up over the mountains,
晴雪为山明。
But the snow melts away when the sun comes around.

(b) A 5-character *quatrains* generated with the user-uploaded image.