Stroke Based Posterior Attention for Online Handwritten Mathematical Expression Recognition

Changjie Wu[1], Qing Wang[1], Jianshu Zhang[1,2], Jun Du[1], Jiaming Wang[1], Jiajia Wu[2], Jinshui Hu[2]

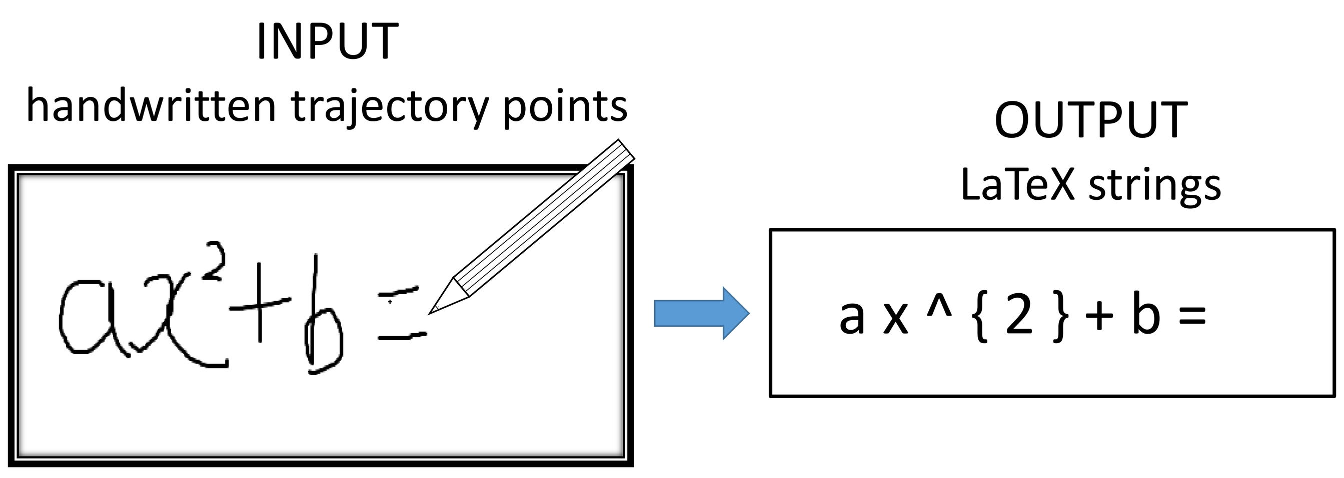
[1]NEL-SLIP, University of Science and Technology of China
[2]iFLYTEK Research, Hefei, China

Introduction

◆ *Background*

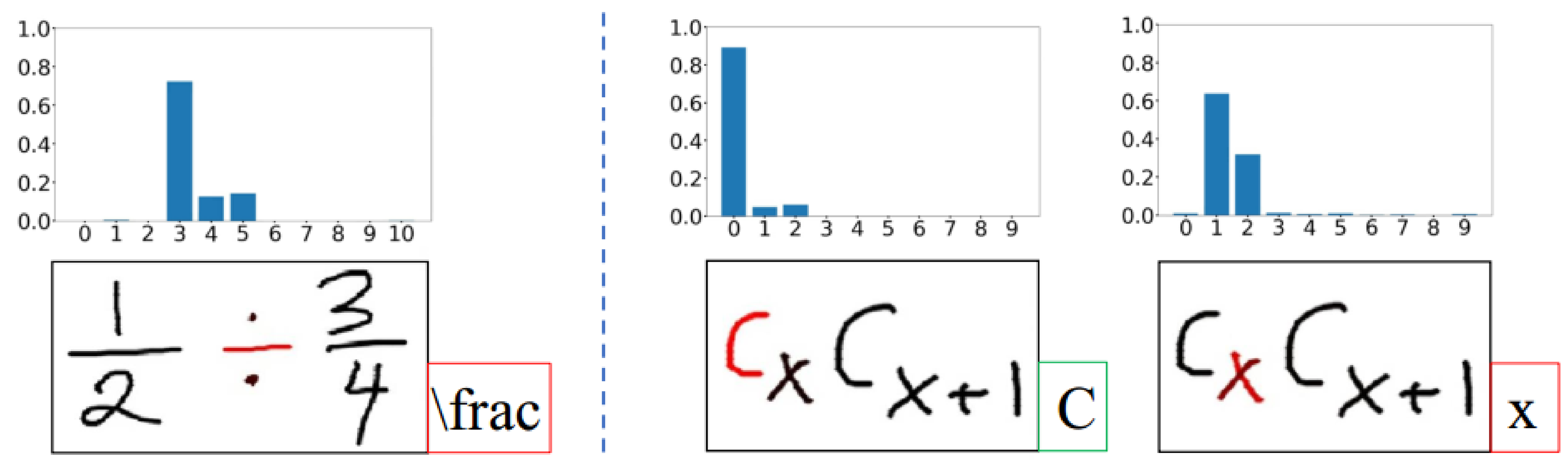
➢ **Online Handwritten Mathematical Expression Recognition (OHMER)**

- **OHMER aims to convert the coordinates of human handwritten trajectory points into a format file that a computer can process such as LaTeX strings and inkml.**



◆ *Challenges*

➢ **Symbol segmentation**
➢ **2D Structural analysis**



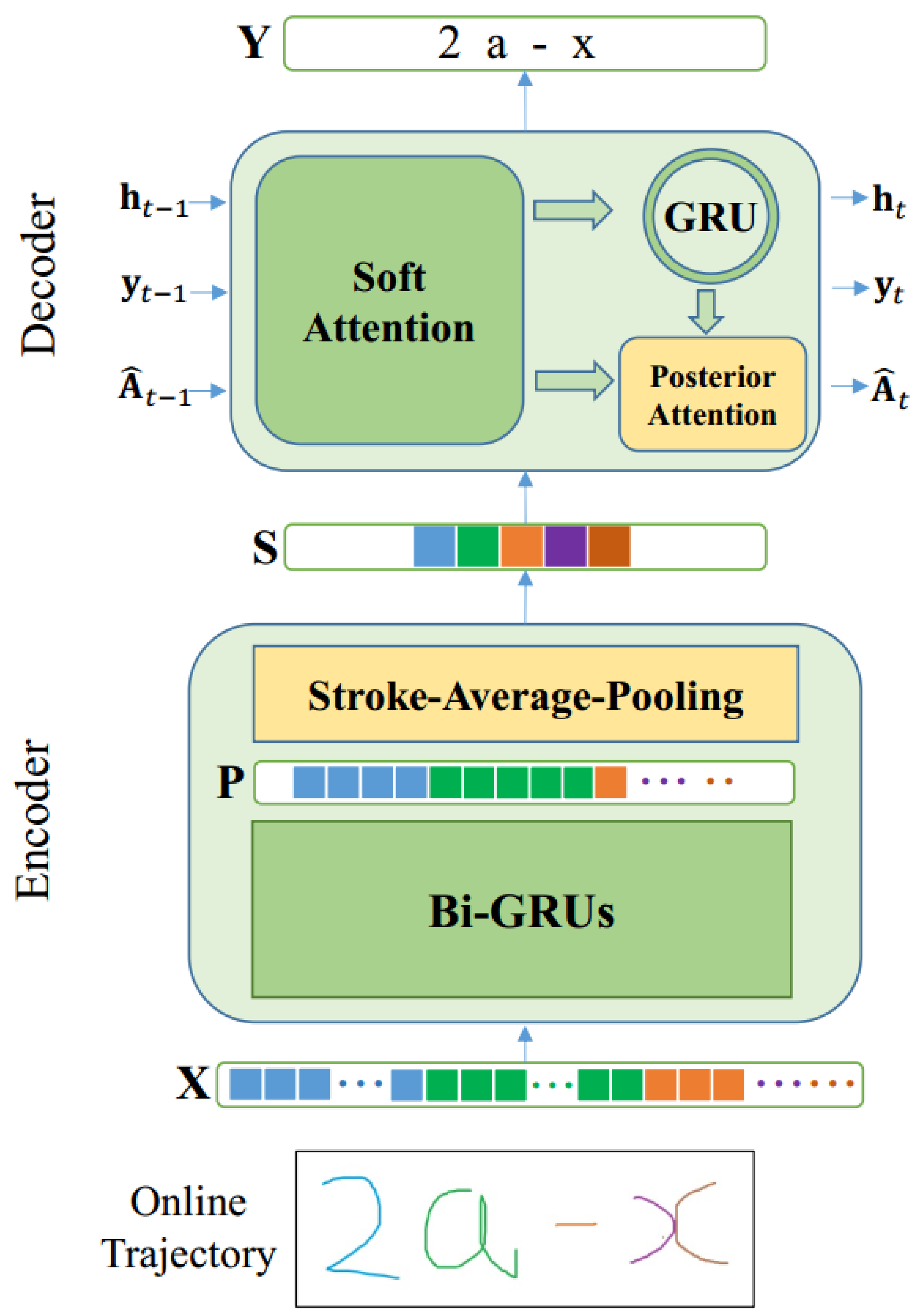
Methods

◆ *overall architecture*

➢ **This is the overall framework of the model: the encoder-decoder framework, which takes the trajectory points as input and outputs a latex string.**

➢ **The stroke average pooling layer aggregate the features from the point level to the stroke level.**

➢ **The posterior attention mechanism is a statistically more reasonable and accurate attention mechanism.**

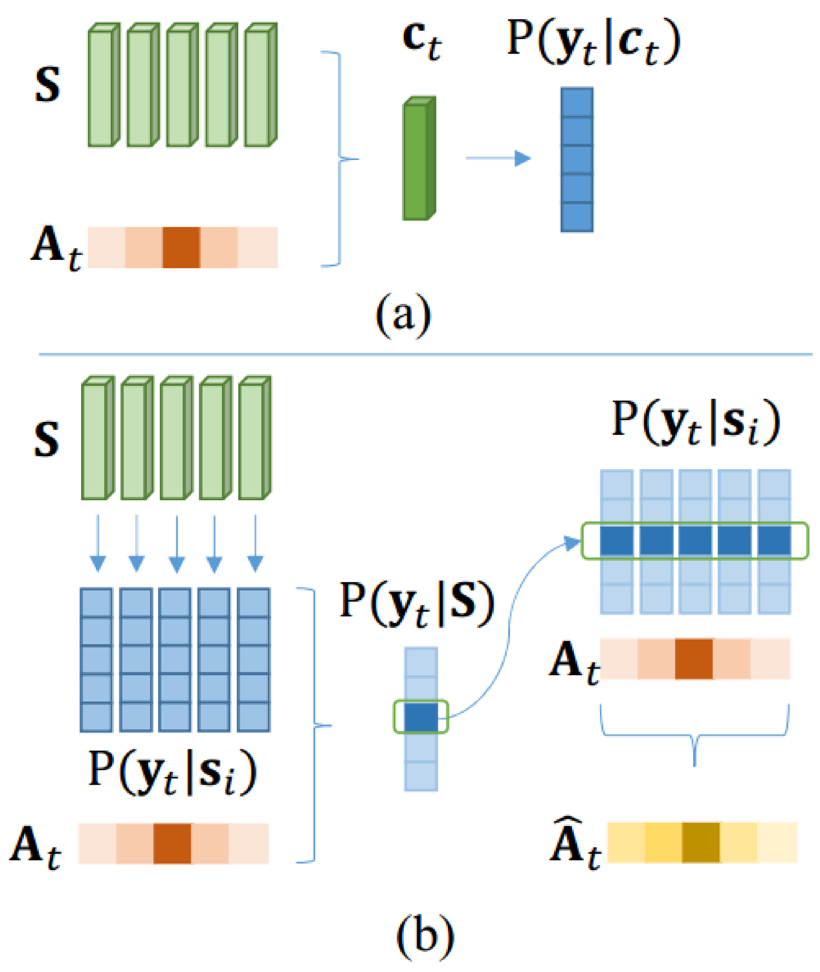


Methods

◆ *Posterior attention*

➢ **The posterior attention are computed by normalizing the soft attention probabilities of all points and taking the output probabilities as the confidence of each point.**

➢ **Posterior attention can get better alignment than soft attention as it considers the posterior information of each point.**



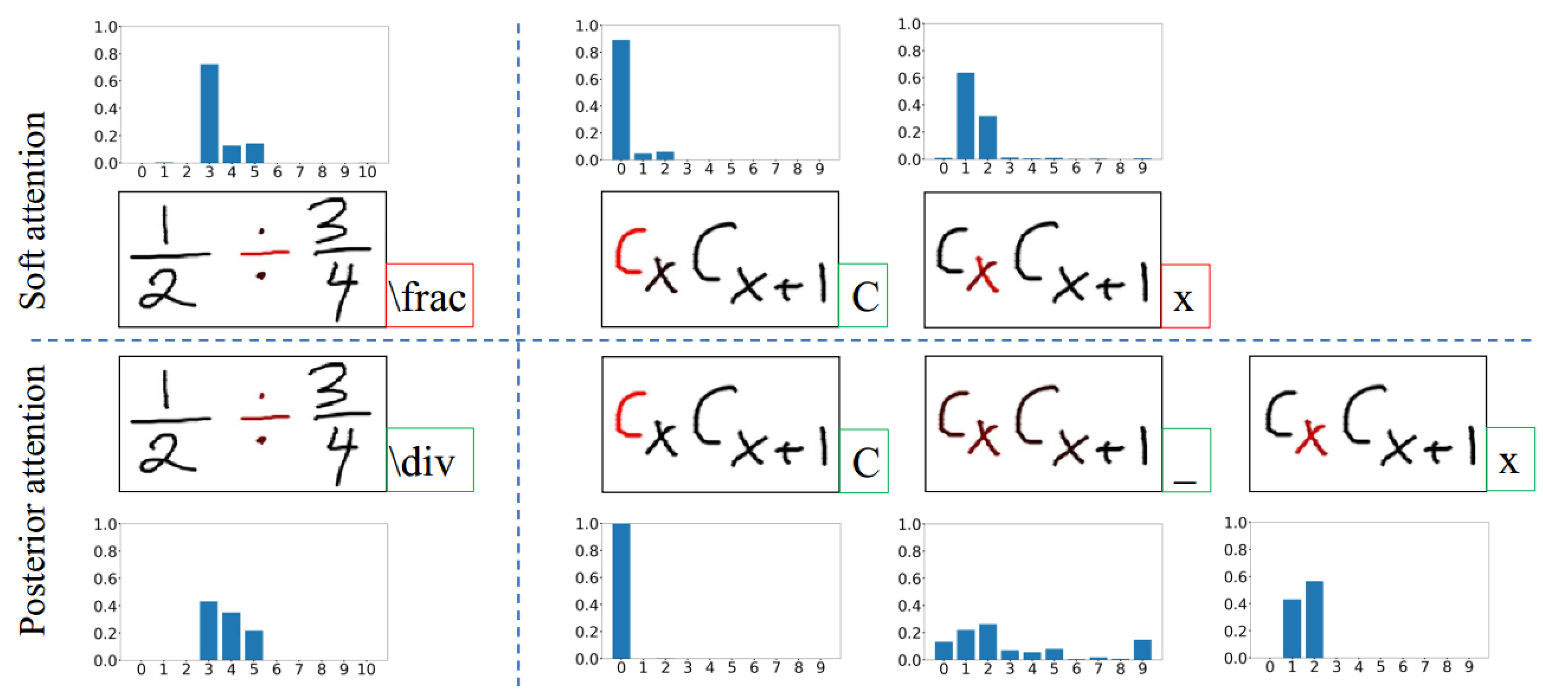
Experiments

◆ *Evaluation of Posterior Attention Mechanism*

TABLE I
COMPARISON OF RECOGNITION PERFORMANCE (IN %) ON CROHME 2014 AND CROHME 2016 BETWEEN SYSTEM I TO IV

System	Attention	Feature Level	CROHME 2014 WER	CROHME 2014 ExpRate	CROHME 2016 WER	CROHME 2016 ExpRate
I	soft	point	13.34	50.71	14.67	45.95
II	posterior	point	11.97	51.28	13.21	47.28
III	soft	stroke	13.29	50.91	14.53	47.60
IV	posterior	stroke	10.44	54.26	12.68	51.75

◆ *Attention visualization*



◆ *Comparison with State-of-the-arts*

TABLE III
COMPARISON OF EXPRATE (IN %) ON CROHME 2014 AND CROHME 2016

System	CROHME 2014 ExpRate	CROHME 2014 ≤1	CROHME 2014 ≤2	CROHME 2014 ≤3	CROHME 2016 ExpRate	CROHME 2016 ≤1	CROHME 2016 ≤2	CROHME 2016 ≤3
Wiris [10]	-	-	-	-	49.61	60.42	64.69	-
Tokyo [10]	-	-	-	-	43.94	50.91	53.70	-
Merge 9 [26]	29.91	39.94	44.96	50.15	27.03	35.48	42.46	-
PGS [27]	48.78	66.13	**73.94**	**79.01**	45.60	62.25	**70.44**	**75.76**
TAP	50.71	65.42	68.73	69.54	45.95	60.77	63.85	64.57
Res-BiRNN [28]	53.35	64.50	70.08	72.92	47.95	60.16	65.56	68.61
Ours	**54.26**	**69.64**	72.65	73.26	**51.75**	**65.18**	68.27	68.99

Conclusion

➢ **The posterior attention mechanism is better than soft attention mechanism**

➢ **The stroke-level feature vectors which contain enough classification information can calculate posterior attention accurately**

➢ **The proposed stroke based posterior attention exhibits higher performance than previous methods.**