

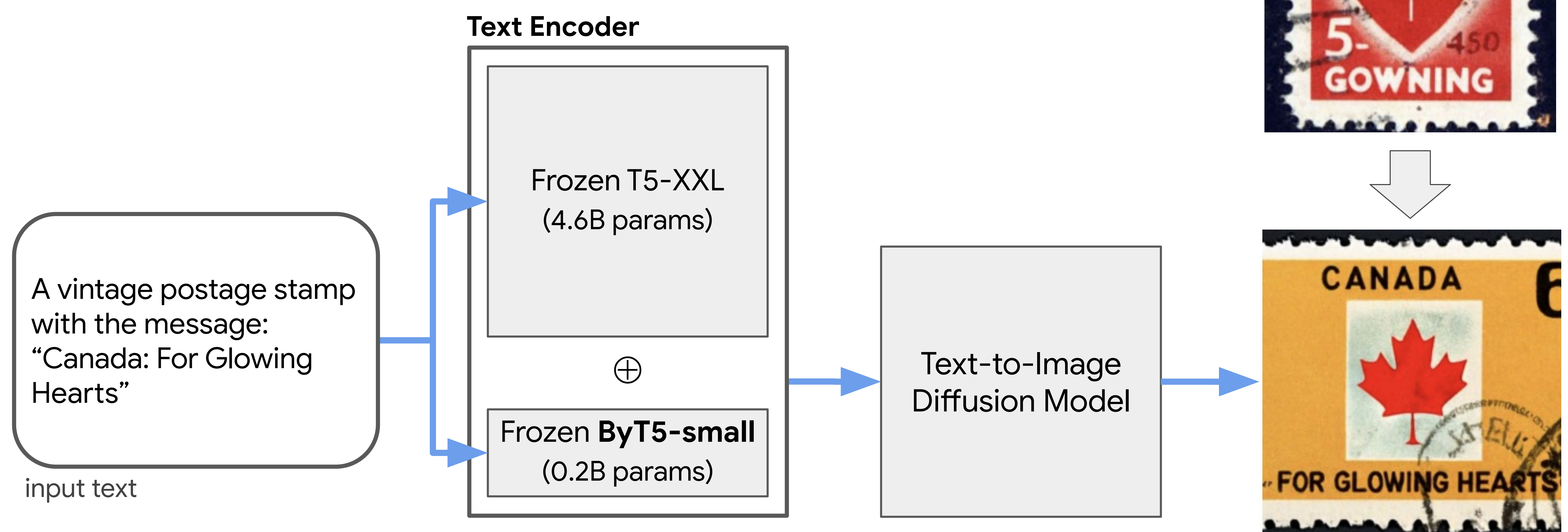
Character-Aware Models Improve Visual Text Rendering

Rosanne Liu*, Dan Garrette*, Chitwan Saharia, William Chan, Adam Roberts, Sharan Narang, Irina Blok, RJ Mical, Mohammad Norouzi, Noah Constant*

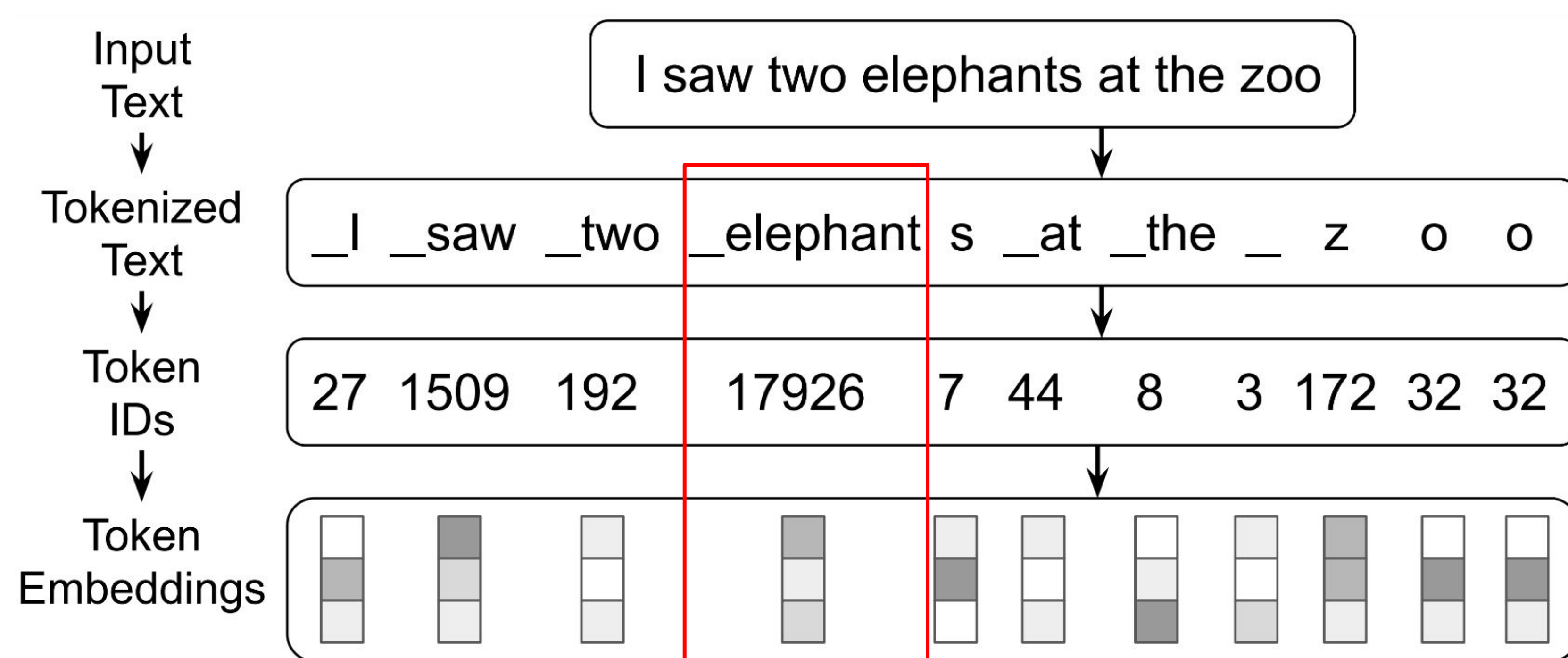


Text-to-image models often fail to correctly render text.
One big reason: Tokenizers hide spelling information.

Our solution: Augment text encoder with character-level inputs.



T5's SentencePiece Tokenization



Example T5 Spelling Failures

Input	Input Tokens	Prediction
taught	_taught	te a ch e d
enthusiasm	_enthusiasm	e n j o y m e n t
crewmembers	_crew member s	crew men
2-up	_2- up	two - up
judge	_judge	justice
subtract	_subtract	addition
temperatures	_temperatures	thermic es
cuisines	_cuisine s	cusinations
Opera	_Opera	Operation
VPN	_VPN	Vicamination

