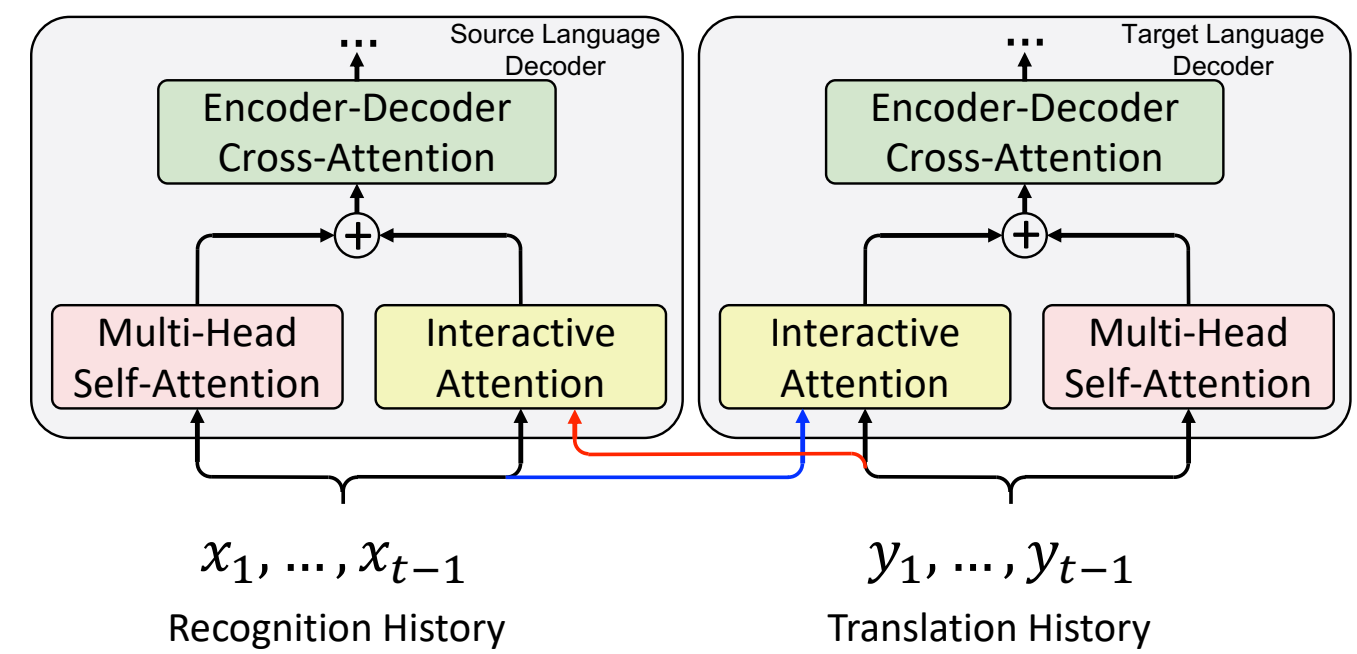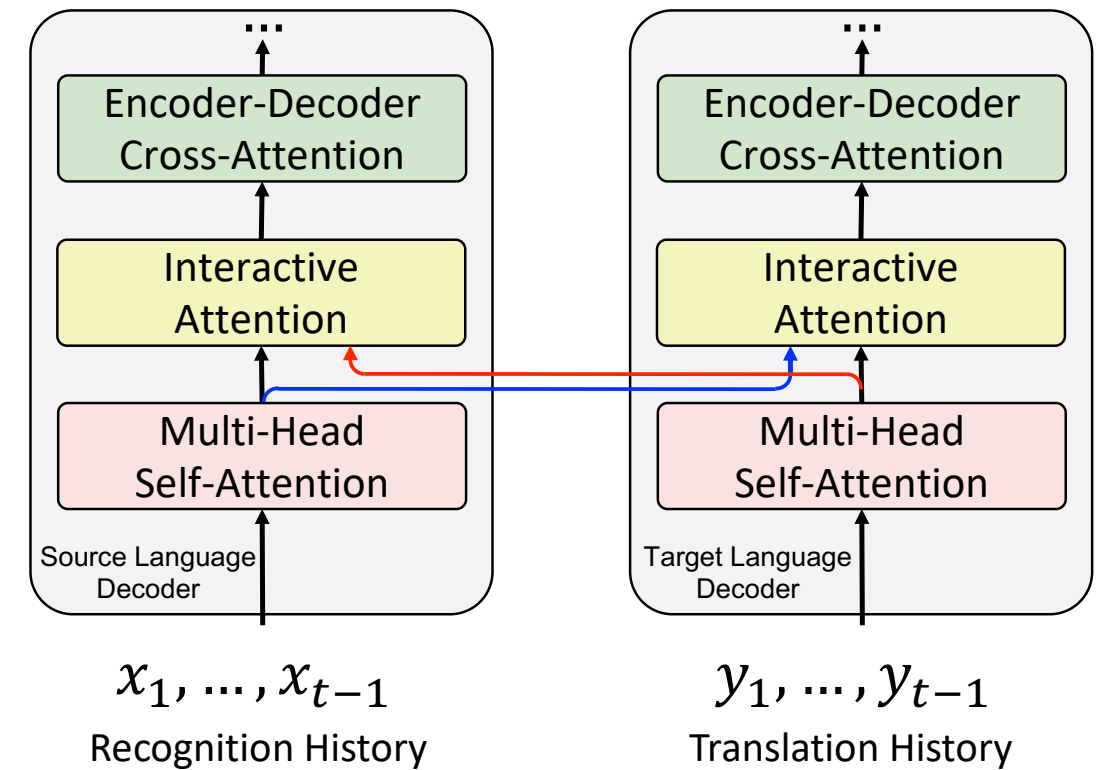(a) Architecture of Cross-modal Cross-lingual Interactive Model for Text Image Translation

(b) Weighted Cross-lingual Interactive Attention

(c) Hierarchical Cross-lingual Interactive Attention