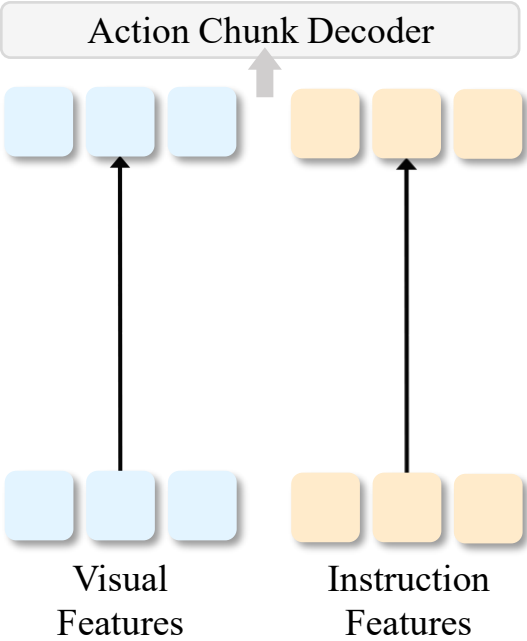
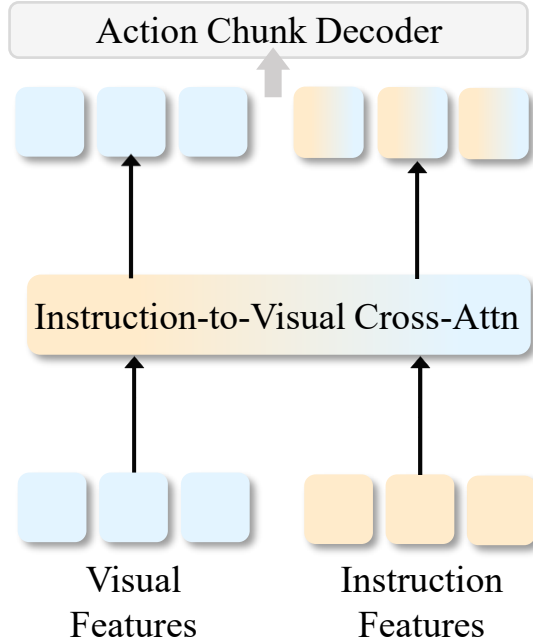
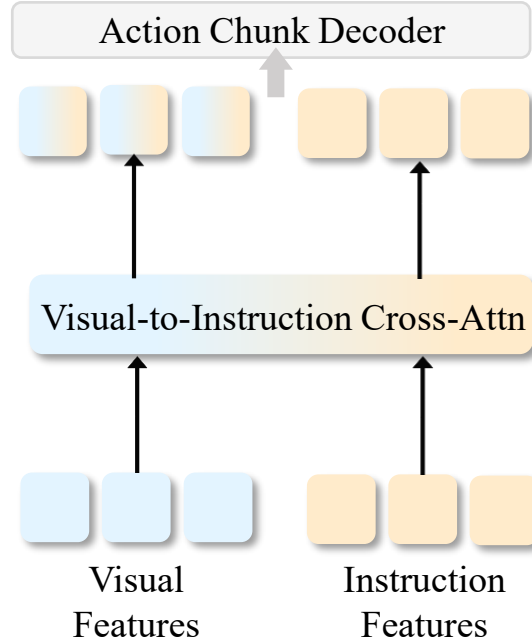




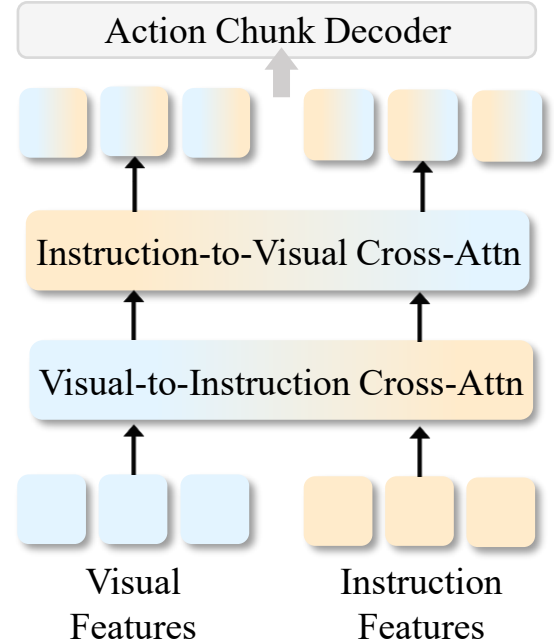
(a) w/o Interaction



(b) Language Queries Visual



(c) Visual Queries Language



(d) Bidirectional Interaction (ours)