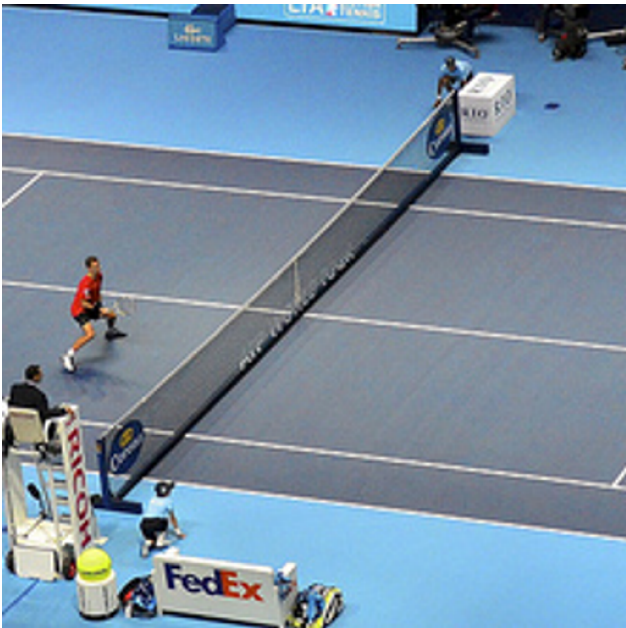
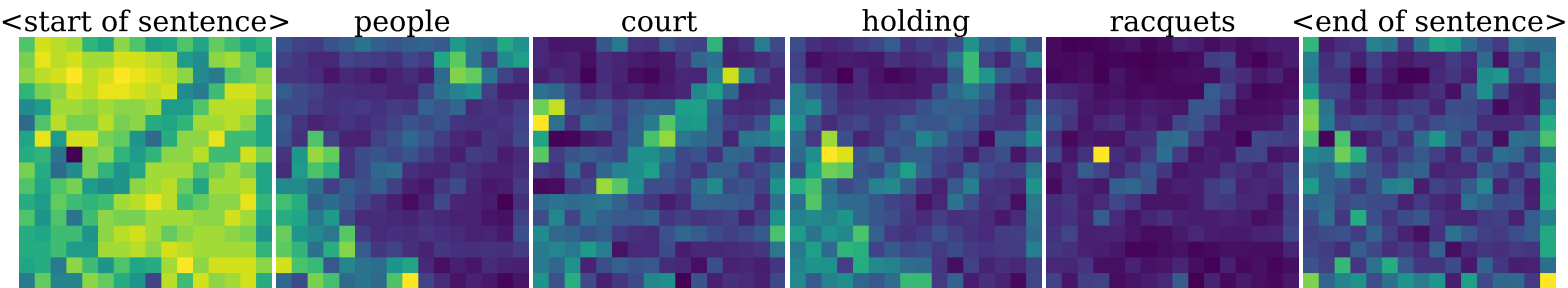


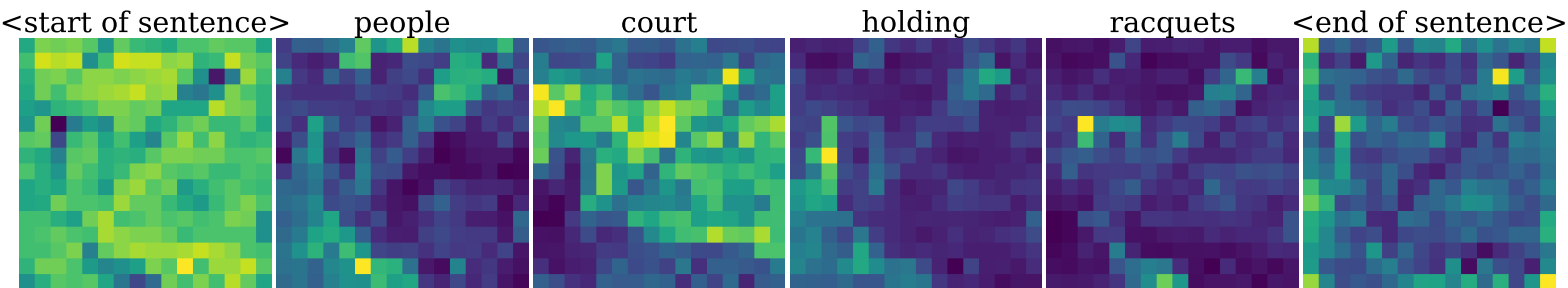
A group of people standing on a court holding racquets



Average Token-Wise Cross-Attention at T=50



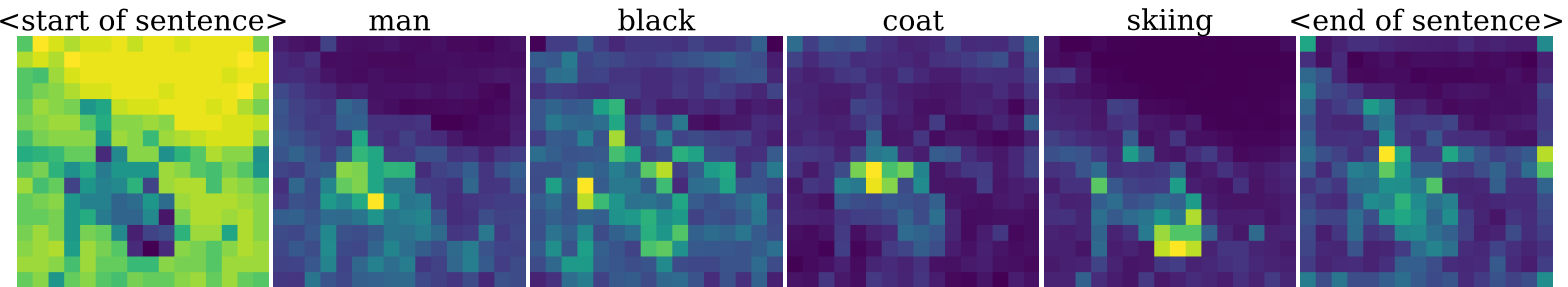
Average Token-Wise Cross-Attention at T=500



a man wearing bright green pants and black coat skiing



Average Token-Wise Cross-Attention at T=50



Average Token-Wise Cross-Attention at T=500

