

Temporal Video Grounding

Zoom in to see mountains

[SEG1]: A camera pans across a mountainous landscape showcasing rugged terrain and varying elevations with patches of lush greenery and winding rivers carving through the valleys.

Temporal Action Localization

Playing badminton

[SEG1]: Two children play badminton outdoors, hitting a shuttlecock back and forth with laughter as a gentle breeze rustles the trees around them.

[SEG2]: A competitive game unfolds as two individuals serve and lunge to return the shuttlecock, their movements swift and precise in the heat of the match.

Video Highlight Detection

Skiing

[SEG1]: A skier in a red jacket descends a snowy slope, skillfully navigating turns as powdery snow sprays up with each sharp carve down the mountain.

Epsodic Memory

What did I put in the plate?

[SEG1]: I saw you prepare food by placing a plate on the countertop and then adding an egg to it.

Video Summarization

Animal

[SEG1]: Affectionate interactions between a person and a cat highlight the bond shared in a domestic setting

[SEG2]: A light-colored cat turns its back to the camera while a striped cat faces right, as a darker cat finds comfort resting against them

Grounded Video QA

What did I stir with?

[SEG1]: The scene shows a a person cooking on a a stove with various utensils and ingredients around hand. The person is using chopsticks to stir the contents of one o the pots, as indicated by the movement of the chopsticks in the frame.