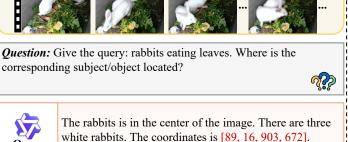
Question: At what point in the video can we observe there is a frisbee away from a dog taking place? Where is the corresponding subject/object located? The event happens in 0.0 - 5.0 seconds. The frisbee is on the left of the dog taking place. The coordinates is [0, 276, 358, 358]. Owen

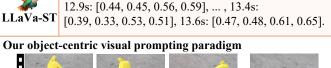
Main Task

Spatial-Temporal Video Grounding



Zero-shot Task

Multi-Object Refering Video Object Segmentation



During the span of 12.9s-13.6s. Object bounding box:





<answer> Target ID: 1, 2. </answer>



<answer> Target ID: 2, Time range: 0.00 to 3.96.

STVG-R1

Qwen

LLaVa-ST



<think> The frisbee (ID2) away from the dog (ID1) is visible on the ground in the first few seconds of the video. The dog is seen interacting with another frisbee. </think>

STVG-R1

</answer>

<think> The rabbit (ID 1) is seen eating leaves in the first frame, and this behavior continues in the subsequent frames as well. The rabbit (ID 2) is also eating leaves. </think>