



Prompt) You will be presented with a video and several pieces of information. One piece of information is consistent with the video while the others are not. Please identify the information that consistent with the video and generate a video caption accordingly.\nInformation A: {'subject': 'entire video', 'order': 'A girl is shown reading a book and a body of water is shown simultaneously'}\nInformation B: {'subject': 'entire video', 'order': 'A girl is shown reading a book, then a body of water is shown'}\nInformation C: {'subject': 'entire video', 'order': 'A body of water is shown, then a girl is shown reading a book'}\n

Ground Truth) B. A girl is shown reading a book, then a body of water is shown

LLaVA-OV-7B) A girl is shown reading a book and a body of water is shown **simultaneously**.

+TCD) A girl is shown reading a book and a body of water is shown **simultaneously**.

+DINO-HEAL) A girl is shown reading a book and a body of water is shown **simultaneously**.

+SEASON) A girl in a black floral dress lies on the grass, **holding an open book in her hands**, deeply engrossed in her **reading**. The serene environment around her includes tall grass and a body of water, with the sun casting a warm, golden light over the scene. As the video progresses, the girl continues to read, with the **body of water becoming more prominent in the background**, enhancing the peaceful and contemplative atmosphere.