## Temporal Video Grounding

- Localize the clip described by the given textual query '{query}' in the video and provide a more detailed description of the clip.
- Identify the clip that semantically matches the given textual query '{query}'
  in the video and provide a description with more relevant temporal and
  spatial details.
- Given a video and a textual query '{query}', find the clip described by the query in the video and describe it with more important details."
- Localize the clip related to the textual query '{query}' in the video and enhance the description by adding more temporal and spatial details relevant to the query.
  - Find the video clip corresponding to the textual query '{query}' and enrich the query with additional but relevant temporal and spatial details.

## **Extractive Video Summarization**

- Summarize the video into a brief synopsis by focusing on the clips that contain the most important visual cues related to the video domain '{query}'.
- Identify the clips in the video that contain the most important visual cues related to the video domain '{query}' and provide a concise synopsis of the video.
- Given a video and a video domain '{query}', summarize the clips that contain the most important visual cues to generate a concise synopsis of the video.
- Summarize the video into a concise synopsis by focusing on the clips with the most important visual cues related to the video domain '{query}'.
- Summarize the video into a concise synopsis by localizing the most important clips with visual cues related to the video domain'{query}'.

## Video Highlight Detection

Find a representative clip in the video that contains the highlight moment depicted by the given query '{query}' and provide a description to help the user understand what is happening in the highlight moment

- Identify a representative highlight clip in the video that corresponds to the query '{query}' and provide a detailed description rich in important visual activities for the user to better understand the moment.
- Given a video and a query '{query}', find a representative highlight clip in the video that corresponds to the query and describe it with necessary details related to the query for better understanding.
- Find representative highlight clip in the video that corresponds to the query '{query}' and enrich the query with more relevant temporal and spatial details to facilitate better understanding.
- Localize a representative highlight clip in the video that corresponds to the query '{query}' and describe it by focusing on highlight-worthy details to enhance the user's experience.

## Temporal Action Localization

- Localize all the clips related to the human action '{query}' and provide a concise description for each of them.
- Identify all the clips that depict the human action '{query}' in the video and provide for each of them a concise description.
- For the given action '{query}', localize all the corresponding clips in the video and describe each of them by focusing on action-related visual activities.
- Extract all the clips in the video that depict the action '{query}' and provide a concise description of each clip.
- Given a video and an action '{query}', find all the clips in the video that depict the action and concisely describe each clip.