Table 1: Comparison of Video Summaries based on Text Prompt Categories

Input Data	Prompt 1 (Concise)	Prompt 2 (Detailed)	Prompt 3 (Absolute Truth)	Prompt 4 (Exact Fields)	Prompt 5 (Follow Exactly)	Prompt 6 (Integrated)	Prompt 7 (Strict Sections)
	Summarize the video briefly (2–3 lines). Avoid details and keep the response concise.	Summarize the video briefly (2–3 lines). Include Details.	Treat the following classification information as ABSOLUTE TRUTH. You are NOT allowed to reinterpret it, modify it, or ignore it. CLASSIFICATION (GROUND TRUTH): {classification_info} You MUST include the following in your answer: - The classification EXACTLY as given. - The fine-grained class EXACTLY as given. - A description of what happens in the video. - Whether the classification matches the video content. Do not omit or alter any required element.	You MUST treat the following classification results as GROUND TRUTH. You MUST extract them EXACTLY and place them in the correct fields below. {classification_info} Extract the following EXACT fields: - 'Model Classification' → goes into the 'Classification' field - 'Fine-grained Class' → goes into the 'Fine-grained Class' field Now output in EXACTLY this format: Classification: <insert Model Classification> Fine-grained Class: <insert Fine-grained Class> Video Description: <describe what happens in the video> Consistency: <state if video matches classification> Do NOT swap the fields. Do NOT rewrite the labels. Do NOT change wording.	Follow the classification information EXACTLY as provided. You may comment on whether the video content agrees with it, but NEVER alter the classification text. Classification Provided: {classification_info} Your output must include: 1. The provided classification as given. 2. A clear description of the video's content. 3. A note on whether the classification and video seem aligned.	Incorporate the following classification into your summary. Do not change the labels, but you may expand on them. Classification: {classification_info} Explain what is happening in the video and indicate whether the visual content supports or contradicts the classification.	You MUST copy the classification information EXACTLY as given, without any changes. {classification_info} Your response MUST contain ONLY the following 3 sections: 1. CLASSIFICATION (copy EXACTLY from above) 2. VIDEO SUMMARY (describe the video) 3. CONSISTENCY (state if video matches classification) Do NOT change section titles. Do NOT rewrite classification. Follow the format EXACTLY.
Model Classification: Normal Activity Fine-grained Class: Haircut	A man getting his beard trimmed with an electric razor in a barber shop.	The video shows a man getting his beard trimmed with an electric razor by a barber in a barber shop. The barber is wearing a white shirt with a decorative bracelet on his wrist. The background features various barber shop equipment and lighting.	Classification: Normal Activity Fine-grained Class: Haircut Description: In the video, a barber is using an electric razor to trim the hair on a man's head and sideburns. The man is seated in a barber chair, and the barber is focused on the task at hand. The background shows various barber tools and equipment, indicating a professional setting. The focus is on the hair cutting process, with the barber's hands and the razor being the main elements of the scene.	Classification: Normal Activity Fine-grained Class: Haircut Video Description: A man is getting his hair cut by a barber in a barber shop. The barber is using a hair clipper to trim the man's hair and beard. The background shows various barber shop equipment and tools. The focus is on the barber's hands and the man's hair and beard. The lighting is bright and the scene is well-lit. Consistency: The video matches the classification results as it shows a barber cutting a man's hair and beard in a barber shop.	The provided classification is "Normal Activity" and the video content seems to align with this classification. The video shows a barber cutting a man's hair and beard with a hair clipper in a barber shop. The scene is well-lit and focused on the hair cutting process, indicating that it is a routine activity.	The video captures a moment in a barbershop where a barber is using an electric razor to trim a man's beard. The scene is set in a dimly lit barbershop with various tools and equipment visible in the background. The barber's hand is visible, holding the razor, and the man's hair and beard are the main focus of the video. The visual content supports the classification of the activity as a haircut, as it shows the barber using a hair cutting tool to trim the man's beard.	1. CLASSIFICATION: Normal Activity 2. VIDEO SUMMARY: The video shows a man getting his hair cut by a barber in a barber shop. The barber is using an electric razor to trim the man's hair and beard. The background is blurred, indicating that the focus is on the man getting his hair cut. 3. CONSISTENCY: The video matches the classification of "Normal Activity" and "Haircut" as it shows a barber cutting a man's hair and beard in a barber shop.
Model Classification: Normal Activity Fine-grained Class: Cricket-Shot	A cricket player swings his bat and hits the ball.	The video shows a cricket player swinging a bat and hitting a ball. The player is wearing a helmet, gloves, and pads. The ball is visible in the air, and the player is captured in mid-air. The video is shot against a black background.	Classification: Normal Activity Fine-grained Class: Cricket-Shot Description: The video captures a cricket player in action, swinging a bat and hitting a ball. The player is wearing protective gear, including a helmet and pads, and is captured in a mid-action shot against a black background.	Classification: Normal Activity Fine-grained Class: Cricket-Shot Video Description: A cricket player is seen swinging a bat and hitting a ball, captured in slow motion against a black background. Consistency: The video matches the classification results.	The provided classification is "Normal Activity" and the fine-grained class is "Cricket". The video content appears to be a cricket player in action, swinging a cricket bat. The player is wearing protective gear, including a helmet and pads, and is captured in a dynamic pose, indicating a moment of action during a cricket match. The classification and video seem aligned as they both depict a cricket player in action.	In the video, a cricket player is captured in mid-action, swinging a cricket bat. The player is wearing a helmet and protective gear, and is positioned against a black background. The player's body is slightly twisted, indicating a powerful swing. The video supports the classification of Normal Activity, as it shows a cricket player in action, engaging in the sport of cricket.	1. CLASSIFICATION: Normal Activity 2. VIDEO SUMMARY: The video captures a cricket player in action, swinging a bat and hitting a ball. The player is wearing protective gear, including a helmet and pads, and is captured in slow motion against a black background. 3. CONSISTENCY: The video matches the classification of Normal Activity and Fine-grained Class: Cricket-Shot.
Model Classification: Normal Activity Fine-grained Class: Apply-Lipstick	A woman applies lipstick while looking at her phone.	The video shows a woman applying lipstick while using her phone. She is standing in a city at night, with a blurred background of lights.	Classification: Normal Activity Fine-grained Class: Apply-Lipstick Description: A woman with curly hair is applying lipstick while holding a smartphone in her other hand. The background is blurred, indicating that the focus is on the woman and her actions.	Classification: Normal Activity Fine-grained Class: Apply-Lipstick Video Description: A woman with curly hair is applying lipstick while holding a phone in her hand. The background is blurred, indicating that the focus is on the woman and her actions.	Classification: Normal Activity Description: The video shows a woman applying lipstick while holding a smartphone. The background is blurred, indicating an outdoor setting at night. Note: The classification aligns with the video content as it captures a common activity of applying lipstick.	In the video, a woman is seen applying lipstick while holding a smartphone. The scene is set in an urban environment at night, with blurred lights in the background. The woman is wearing a checkered blazer over a white shirt and has curly hair. She is using a pink lipstick and appears to be focused on her makeup application. The video captures the woman's detailed attention to her lips, indicating that she is engaged in a routine activity. The classification of "Normal Activity" is supported by the woman's focused and deliberate actions, as well as the context of the urban setting.	1. CLASSIFICATION: Normal Activity 2. VIDEO SUMMARY: The video shows a woman applying lipstick while holding a smartphone in her other hand. The background features a blurred cityscape with lights. 3. CONSISTENCY: The video matches the classification of "Normal Activity" as it captures a common everyday activity of applying lipstick.
						Continued on next page	

Table 1 – continued from previous page

Input Data	Prompt 1	Prompt 2	Prompt 3	Prompt 4	Prompt 5	Prompt 6	
Model Classification: Normal Activity Fine-grained Class: Brushing-Teeth	A woman brushes her teeth in front of a mirror.	The video shows a woman brushing her teeth in front of a mirror. She is seen holding a toothbrush and brushing her teeth.	The classification EXACTLY as given is "Normal Activity". The fine-grained class EXACTLY as given is "Brushing-Teeth". The video content shows a woman brushing her teeth in front of a mirror.	Classification: Normal Activity Fine-grained Class: Brushing-Teeth Video Description: A woman is seen brushing her teeth in front of a mirror.	The provided classification is correct. The video shows a woman brushing her teeth in front of a mirror. The fine-grained classification of "Brushing-Teeth" aligns with the activity depicted in the video. The woman is actively engaging in the act of brushing her teeth, which is a normal and essential part of oral hygiene.	In the video, a woman is seen brushing her teeth in front of a mirror. She is using a purple toothbrush and appears to be in the process of cleaning her teeth. The visual content supports the classification of "Normal Activity" and "Brushing-Teeth" as it shows a common daily routine.	1. CLASSIFICATION: Normal Activity 2. VIDEO SUMMARY: The video shows a woman brushing her teeth in front of a mirror. She is seen holding a toothbrush and brushing her teeth carefully. 3. CONSISTENCY: The video matches the classification of "Normal Activity" and "Brushing-Teeth" as it depicts a woman performing a common dental hygiene routine.
Model Classification: Normal Activity Fine-grained Class: Baby-Crawling	A baby crawls on its belly.	The video shows a baby crawling on a white floor. The baby is wearing a diaper and appears to be happy and content. The camera captures the baby from various angles, including a close-up shot of its face. The baby eventually stands up and walks away.	The classification EXACTLY as given is "Normal Activity". The fine-grained class EXACTLY as given is "Baby-Crawling". In the video, a baby wearing a diaper crawls on all fours across a white surface. The baby appears to be moving with a gentle, rhythmic motion, and its face is turned towards the camera, showing a sense of curiosity and interest.	Classification: Normal Activity Fine-grained Class: Baby-Crawling Video Description: In the video, a baby is seen crawling on a white surface. The baby is wearing a diaper and appears to be moving around on all fours. The video captures the baby's movements as it explores its surroundings.	The provided classification is "Normal Activity" and the fine-grained class is "Baby-Crawling". The video content shows a baby crawling on a white surface. The baby is wearing a diaper and appears to be moving around on all fours. The classification and video seem aligned as the baby is engaged in a typical activity of crawling, which is a common behavior for babies.	In the video, a baby is seen crawling on a white surface. The baby is wearing a diaper and appears to be moving around on all fours. The visual content supports the classification of "Normal Activity" and "Baby-Crawling" as the baby is actively engaging in these activities.	CLASSIFICATION: Normal Activity VIDEO SUMMARY: The video shows a baby crawling on a white surface. The baby is wearing a diaper and appears to be happy and smiling as it moves along the floor. CONSISTENCY: The video matches the classification of "Normal Activity" and "Baby-Crawling" as it depicts a baby crawling on a white surface.