

Prompt: "Generate a concise caption for the input video. The caption should describe the main event, key objects, location context, actions. Focus on factual and clear language, avoiding speculation or overly creative phrasing. "

[Landslide]



References: 1. dirt landslide next to the road with cars surrounded by trees. 2. cars are surrounded by trees and there is a dirt slide. 3. cars are surrounded by trees after a dirt slide. 4. landslide of dirt beside the road with cars surrounded by trees. 5. muddy landslide next to the road with cars surrounded by trees.

LongVA-7B-TPO: Aerial view of a steep hillside covered in lush greenery, with power lines running through the forest.

[Ploughing]



References: 1. flock of birds follows five cultivator machines as they plow a large agricultural land. 2. flock of birds follows five tillers to till a large field. 3. flock of birds follows five cultivators plowing a large farmland. 4. flock of birds follow five machines as they plow a large area of land. 5. flock of birds follow the machines that plow the land.

LongVA-7B-TPO: A large field with a long row of red tractors in a line, all working together to plough (Missing details).