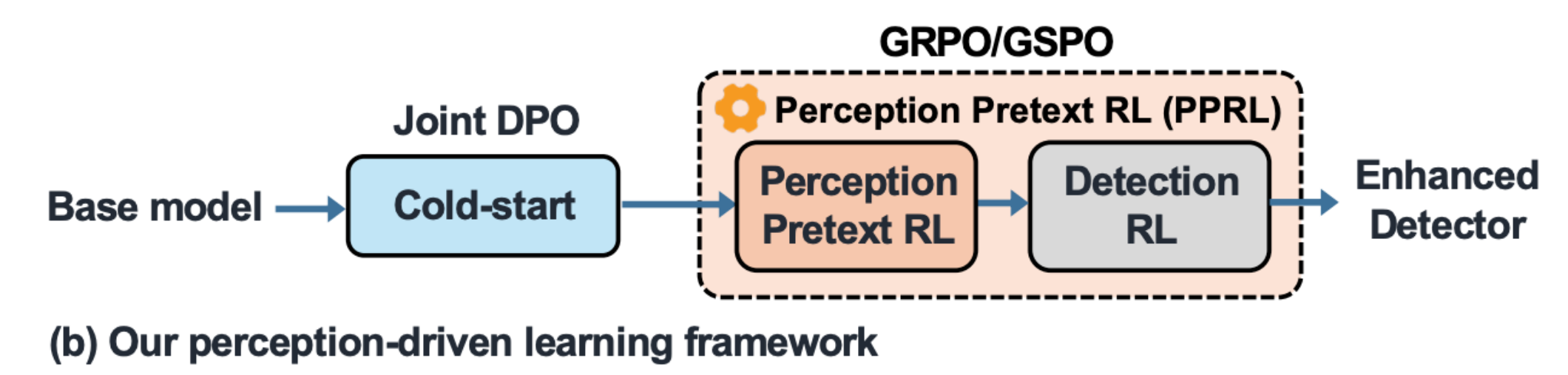
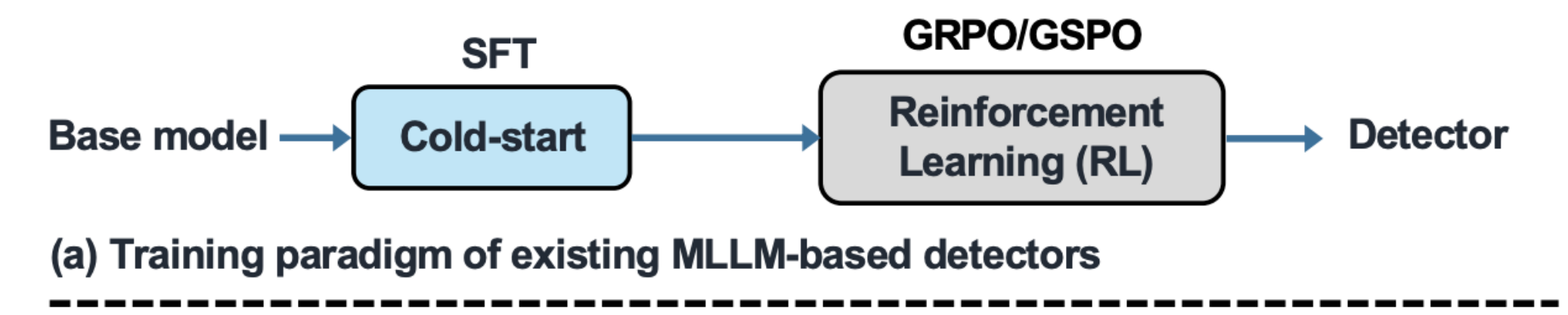
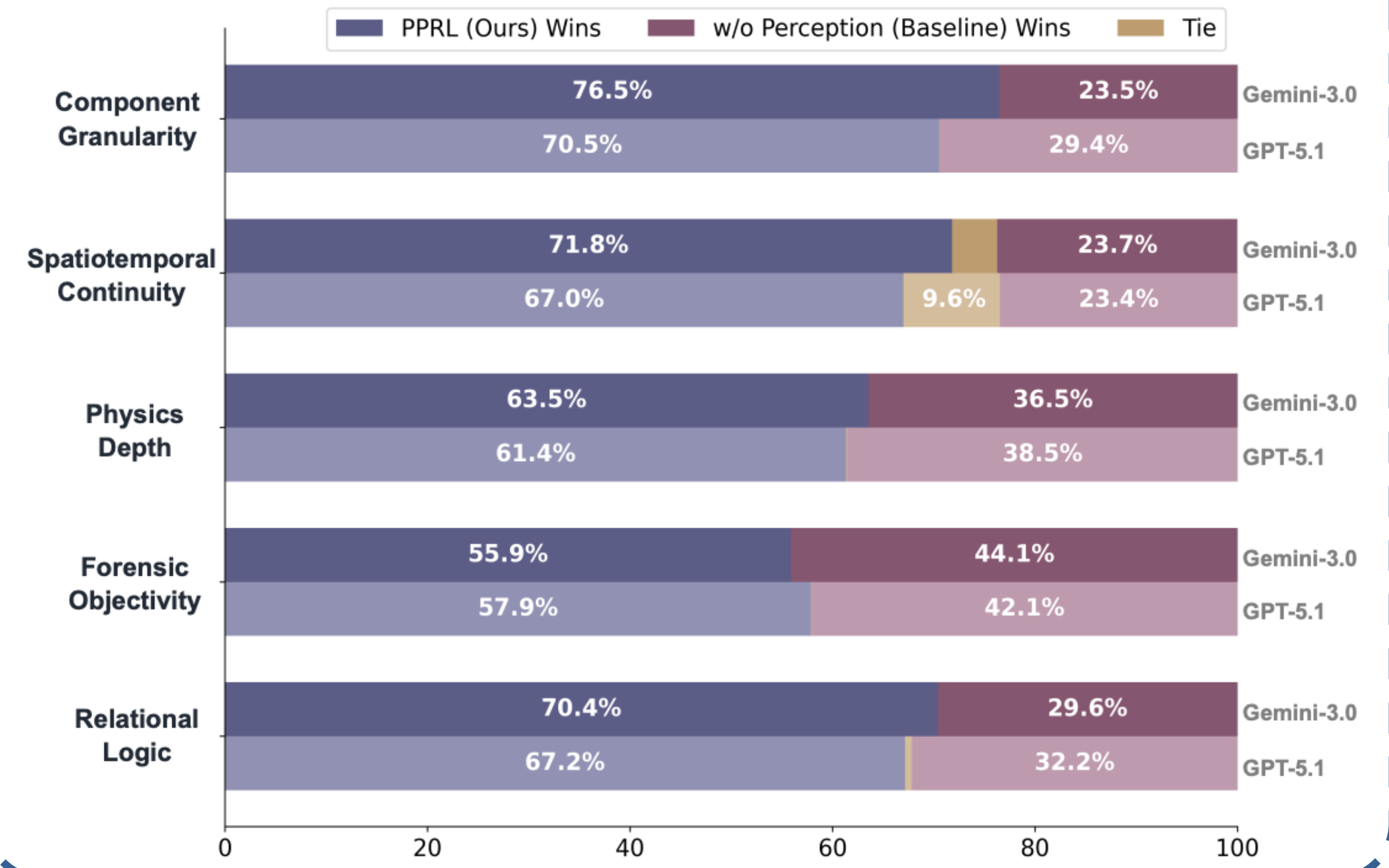


## Motivation & Solution

- **Motivation:**
  - Strong reasoning vs. Poor perception
  - Mechanical analysis after SFT
- **Solution:**
  - Perception-oriented Learning

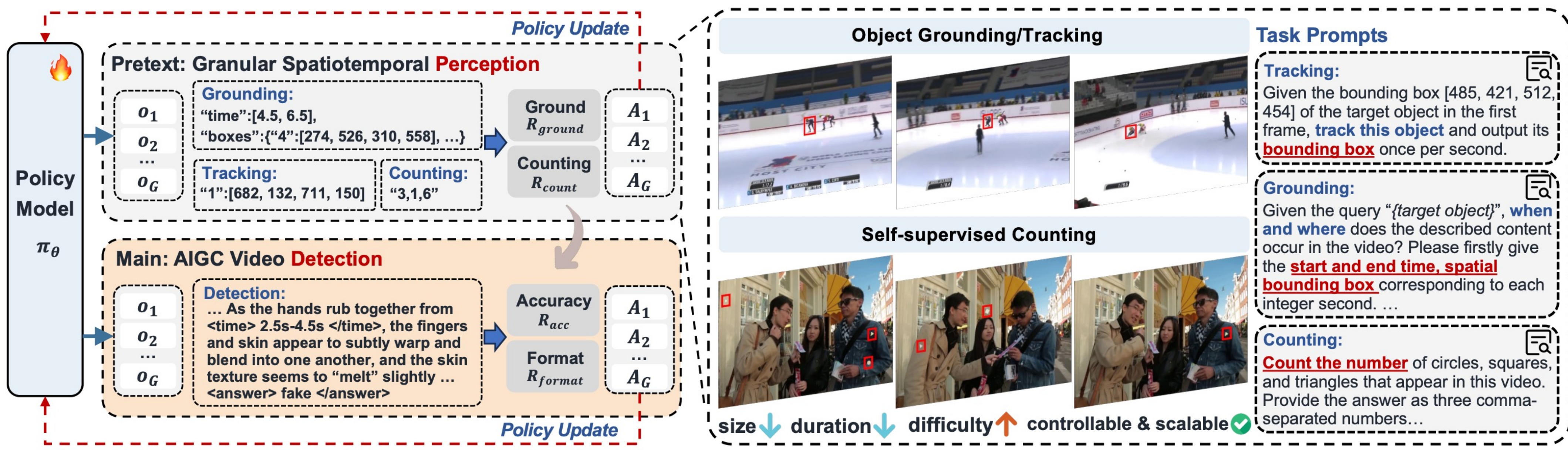


### Why this works?



## VideoVeritas Model

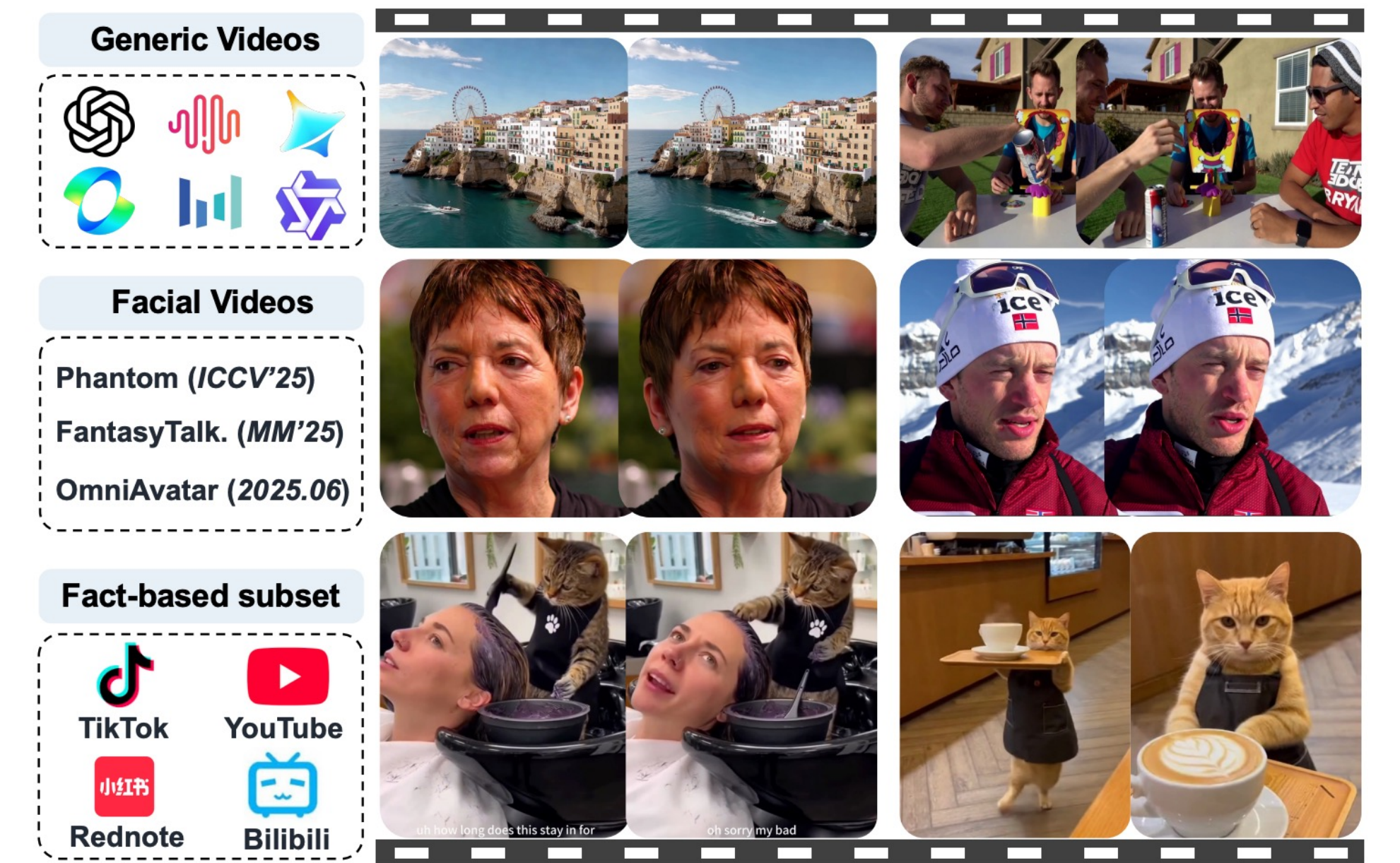
### Core Solution: Perception Pretext RL (PPRL)



- Idea: Perception is idea taken as a foundational phase to detection
- Task: General grounding/tracking & Self-supervised counting

## MintVid Dataset

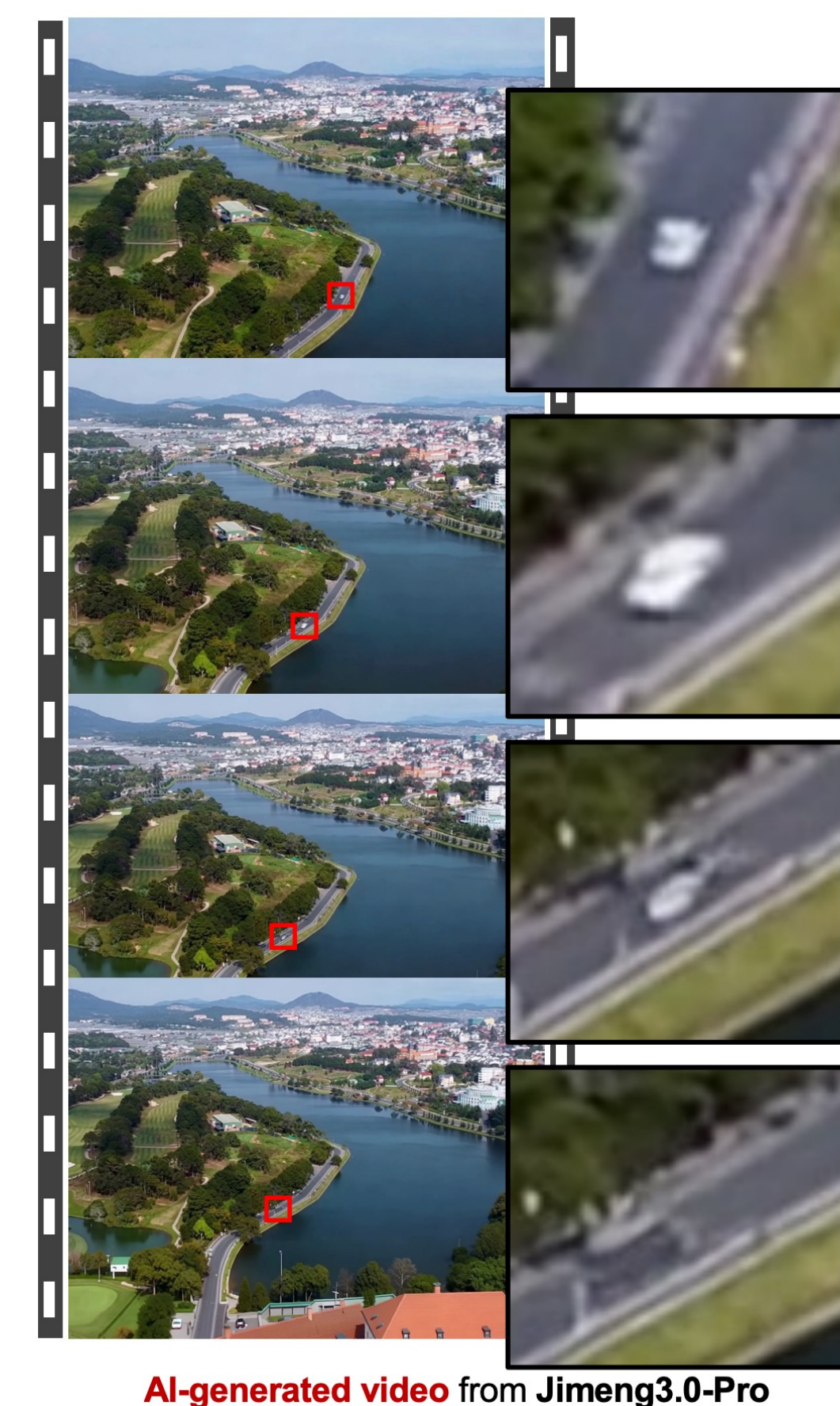
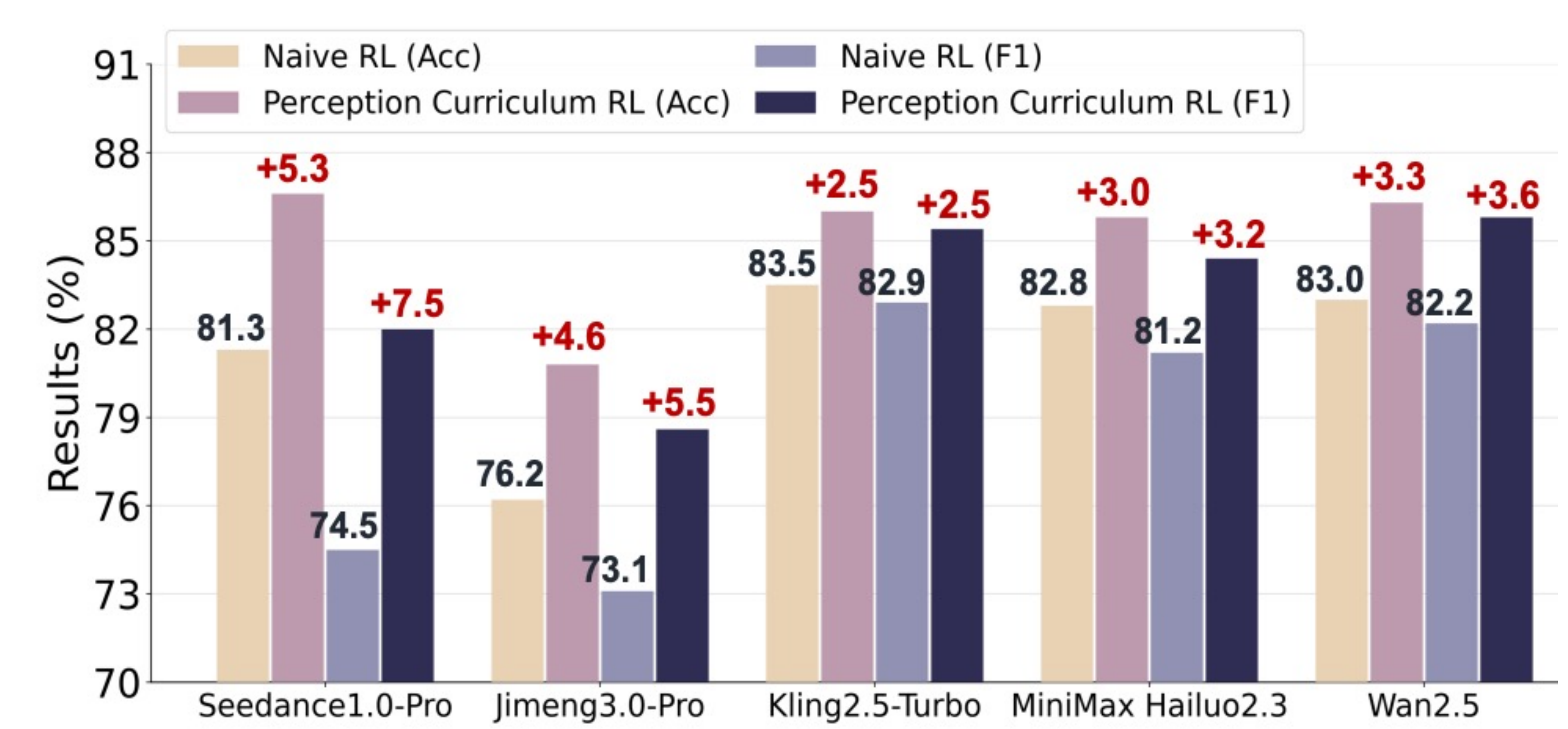
### High-quality & Lightweight



- 3 folds: Generic, Facial, Fact-based
- Sora2, Jimeng3.0Pro, Kling2.5, ...

## Experiments

### Improvements of PPRL



**BusterX++** ✓

<think> Let's analyze this video step by step to determine if it is real or fake:

- Architectural Style**: ... The symmetry and the perfect alignment of the rooftops and windows suggest a high level of digital manipulation, which is not typical of real-world photography or video footage.
- Landscape and Environment**: ... Additionally, the lighting and shadows are very even and lack the subtle variations that would be present in a real-world setting. ...

<answer>B</answer>

**Skyra-RL** ✗

<think> I am conducting a thorough analysis of this aerial video to verify its authenticity. ... I am examining the region <[0.0, 0.54]> at <[0.6, 0.5, 0.7, 0.6]>. In this part of the video, the scene consists of water, a road, and distant hills. There are no people or faces present, and therefore no facial features can be distorted. ....

<answer>Real</answer>

**VideoVeritas (ours)** ✓

... I'll start by scrutinizing the dynamic elements, specifically the vehicles on the road, as they are difficult to render perfectly. Then, I'll examine the textures of the water and the architectural details of the buildings for any signs of instability or unnatural perfection. The scene's realism quickly unravels upon closer inspection. The most immediate anomaly is the behavior of the vehicles on the winding road in the foreground. Throughout the video, from <time> 0.0s-5.0s </time>, cars move along the road, but their appearance is fundamentally flawed. For instance, a light-colored car appears around <time> 1.0s </time>, moves smoothly for a moment, and then abruptly and comically morphs into a different, more angular vehicle shape before disappearing by <time> 1.3s </time>. This is not a camera effect or a reflection; it's a clear failure of the model to maintain the object's consistent form and structure, a hallmark of object permanence failure. ....

<answer> fake </answer>