



Please describe the image in detail.



LVLM

Initial Response: The image depicts a wide ...
There are several traffic cones placed along the road ... The traffic cones ... **In addition to the traffic cones, there are a few cars on the road, with one car positioned closer to the left side of the image ...**
The presence of these cars suggests that the road ...



Critical-object identification

RAM++

Grounding-DINO-B



<person>,
<motorcycle>,
...

Critical-object description

BLIP-2-
FlanT5_{XXL}

Entity cross-checking results

a highway with a bridge over it

a person is riding a motorcycle on a road

...

Critical-Object Enhancement Framework



The image depicts a wide ... There are several traffic cones placed along the road ... The traffic cones ... **In addition to the traffic cones, there are a few cars on the road, with one car positioned closer to the left side of the image ...**
The presence of these cars suggests that the road ...

Some additional information includes: a highway with a bridge over it, a person ...

Refined Image-Text Pair



The image depicts a wide, empty road ...
There are several traffic cones placed ...
The traffic cones are scattered in various ...
In addition to ..., there are a few cars ...
The presence of these cars suggests that ...

Sentence split & key entity extraction

Llama-3.1

road.bridge
traffic cone
traffic cone
car
car



Entity cross-checking

BLIP-2-
FlanT5_{XXL}

InstructBLIP-
FlanT5_{XXL}

InstructBLIP-
Vicuna-13B

entity
✓ ✗

Is there a road in the image? Is there a bridge ...
Is there a traffic cone ...?
Is there a car ...?
...



road.bridge; traffic cone; ...



Hallucinatory text correction

Llama-3.1

Initial response

The image depicts a wide, empty road with a ... **(No change)**
There are several traffic cones placed along the... **(No change)**
The traffic cones are scattered in various positions... **(No change)**
In addition to the traffic cones, there are a few cars on the road... **(Remove hallucination)**
The presence of these cars suggests that the road is not entirely ... **(Remove hallucination)**



Hallucination Cross-Checking Framework