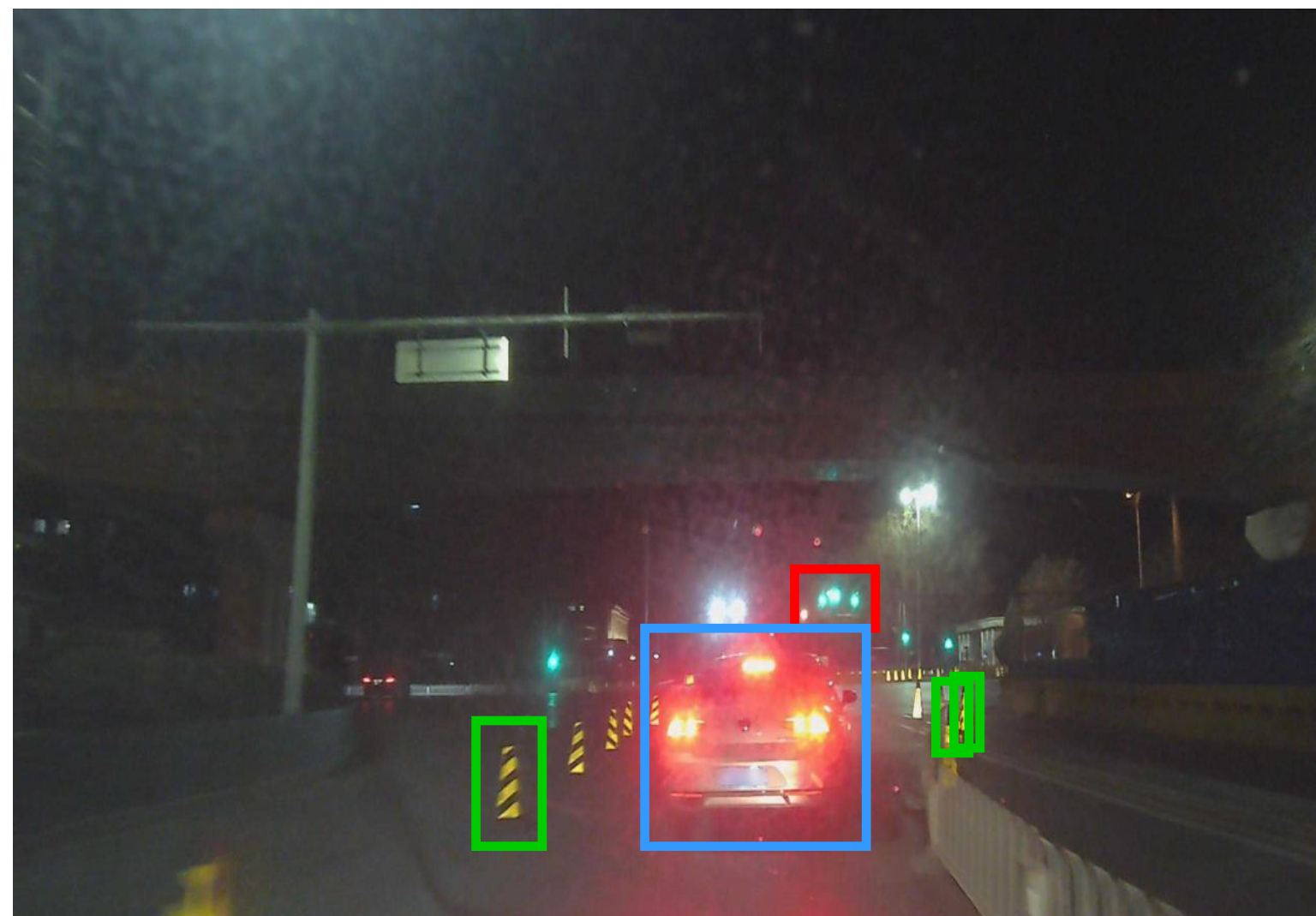


Describe the image in detail.



**LLaVA-1.5-7B:** The image depicts a busy highway at night, with **multiple cars** driving under a bridge. There are **at least nine cars** visible in the scene, with some cars positioned closer to the foreground and others further back. The cars are moving in different directions ...

**+Ours:** The image depicts a highway at night, with **a car** driving under a bridge. **There is at least one car visible in the scene.** **The cars are moving in the same direction, ...**

(a) "Scene\_graph": {

"objects": [

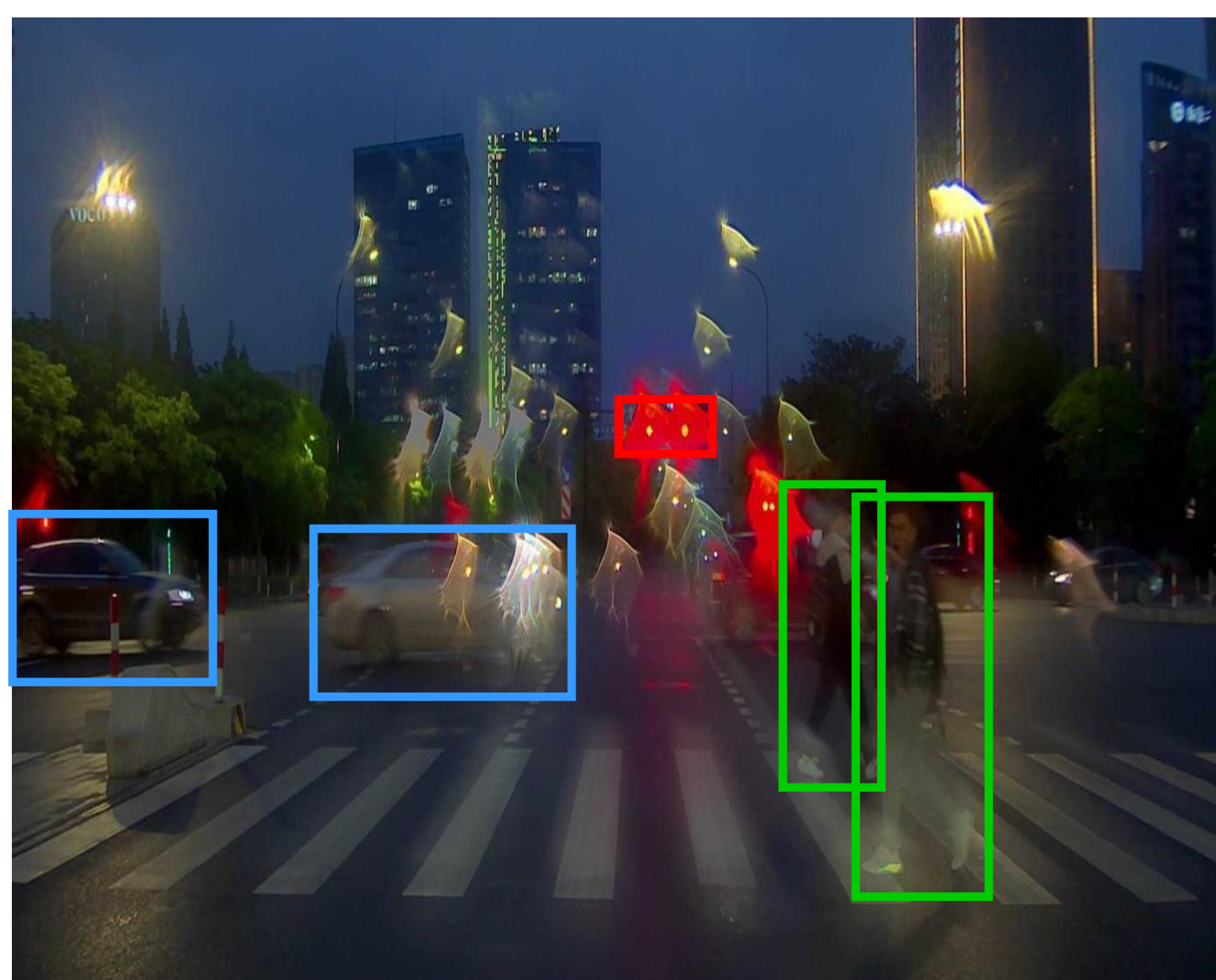
{"type": "**traffic light**", "position": "above the intersection", "status": "green", "boxes": [0.52, 0.61, 0.63, 0.68]},

{"type": "**vehicle**", "position": "in front on the road", "status": "stationary", "boxes": [0.48, 0.69, 0.64, 0.91]},

{"type": "**traffic cone**", "position": "on the side of the road", "status": "visible", "boxes": [[0.72, 0.74, 0.73, 0.79], [0.68, 0.73, 0.69, 0.78], [0.37, 0.80, 0.39, 0.89]]},

"relationships": [{"source": "traffic\_light", "target": "vehicle", "relation": "controlling"}, {"source": "vehicle", "target": "traffic\_cone", "relation": "next\_to"}]}

Describe the image in detail.



**LLaVA-1.5-7B:** The image captures a busy city street at night ... The scene is filled with various vehicles, **including cars and a truck**, driving and parked along the street. ... The street is bustling with activity, as several people can be seen walking or standing around, **some of them carrying handbags.**

**+Ours:** The image captures a busy city street at night ... The scene is filled with various vehicles, **including cars**, driving and parked along the street. ... The street is bustling with activity, as several people can be seen walking or standing around.

(b) "Scene\_graph": {

"objects": [

{"type": "**traffic light**", "position": "above the intersection", "status": "red", "boxes": [0.52, 0.39, 0.63, 0.43]},

{"type": "**vehicle**", "position": "in the distance on the road", "status": "stationary", "boxes": [[0.00, 0.53, 0.16, 0.66], [0.25, 0.53, 0.42, 0.67]]},

{"type": "**pedestrian**", "position": "crossing the street", "status": "walking", "boxes": [[0.70, 0.49, 0.79, 0.89], [0.65, 0.50, 0.72, 0.79]]},

"relationships": [{"source": "traffic\_light", "target": "vehicle", "relation": "controlling"}, {"source": "traffic\_light", "target": "pedestrian", "relation": "controlling"}]}