Tier 1
Relational Geometry

Relative Distance:

Choose from food truck(id:1), fire extinguisher(id:5), road sign(id:6), and leafless tree(id:7): which one is the closest to green umbrella(id:4)?

Answer: food truck(id:1)

Relative Direction:

If I am standing by food truck(id:1) and facing fire extinguisher(id:5), is green umbrella(id:4) to my front-left, front-right, back-left, or back-right?

Answer: front-right

Qualitative Ego-Motion:

Which one best describes the camera's overall movement throughout the entire video? Choose from 'straight', 'left turn', 'right turn' or 'U-turn'.

Answer: straight

Tier 2
Static Metric

Object Localization:

At 8 seconds into the video, what is the Euclidean distance of the food truck(id:3) from the camera in meters?

Answer: 21

Absolute Distance:

What's the distance between fire extinguisher(id:5) and leafless tree(id:7) in meters?

Answer: 17

Depth-Aware-Counting:

At 25 seconds into the video, how many traffic cones are visible within 10 meters from the camera?

Answer: 2 (see circles in image)

Tier 3
Dynamic Metric

Absolute Displacement:

What is the displacement distance of person with a white hat and backpack(id:2) between 6s and 9s of the video in meters?

Answer: 6.7

Absolute Speed:

What is the average speed of child riding a bike(id:8) between 25s and 27s of the video in m/s?

Answer: 1.7

Quantitative Ego-Motion:

How long has the camera travelled throughout the entire duration of the video in meters?

Answer: 45