









Timestamp:30.1

HolisticOCR

Q:Please provide all text in video.

GT:[EXITS] [IOE-W][495] [Long] [is] [Expwy] [Eastern] [Long] [is] [Manhattan] [RIGHT] [LANE] [SPEED] [LIMIT] [40]

HolisticReasoning

Q: Based on the video, which of the following statements accurately describe the video?

GT:1. The video was recorded on Interstate 495 (Long Island Expressway) in New York State.

- 2. Vehicles in the right can take Exit 10E.
- 3. The speed limit on this stretch of road is 40 mph.



LocalOCR









Timestamp:2307.6

62s

225

Q:What texts appear between 22s and 52s

GT:[HAPPY] [NEW] [YEAR!!!!] [HELLO] [2024]

LocalReasoning

Q:What is the author's training plan between 52 s and 62 s in the video?

GT: The training plan is divided into push, pull, legs, and full-body parts.



in the video?









457.35



Timestamp:598.7 462.15

525

TextLocalization

Q:At what time in the video does the score "105:83" appear?

GT:[457.3s,462.1s]

TemporalCausalReasoning

Q:When the score is 105:83, what move did Lakers No.23 make?

GT: He stole the ball and dunked it.











90.15





116.39

TextTracking

Q:What are the bounding box of Hero T10ner at the start and at the end of the match interval from 5:59 to 6:15?

GT:[0.46,0.67...], [0.58, 0.69...]

SpatialReasoning

Q:At 5:59 in the match, in which direction is BLG Bin relative to T10ner, and what happens next?

GT:T1 Oner, positioned to the bottom-right of BLG BIN, and T1 Oner attempts to help his teammate gank BLG BIN.