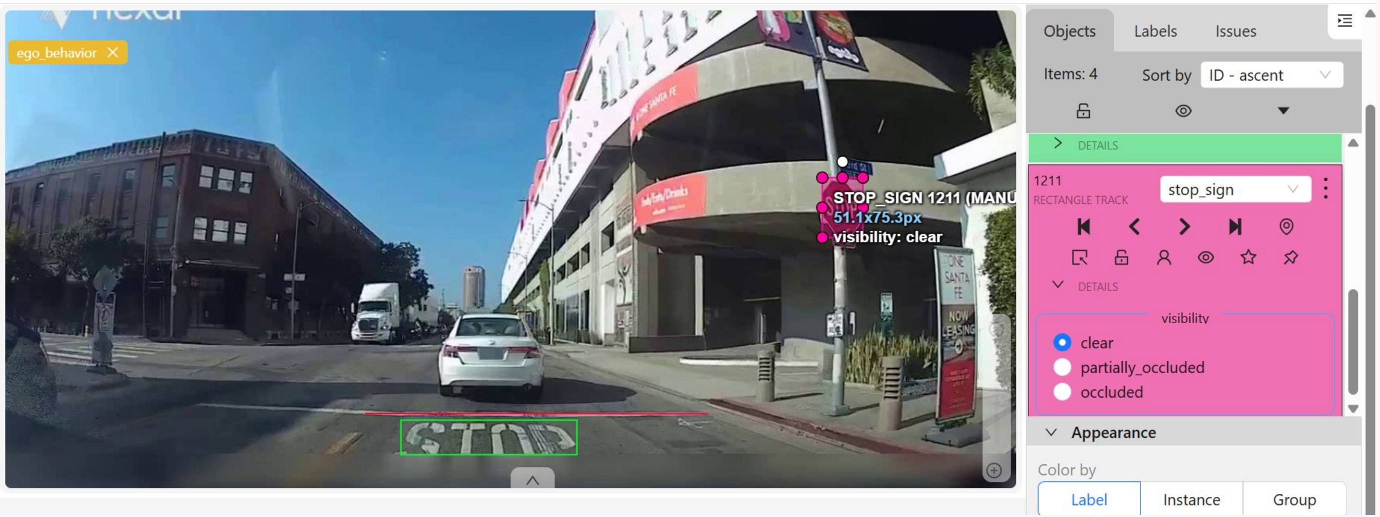
Amit Tal, June 2025

Annotation Protocol
Stop Sign Behavior Analysis



STOP_SIGN 1211 (MANU
51.1x75.3px
visibility: clear

Items: 4

Sort by ID - ascent

1211
RECTANGLE TRACK
stop_sign

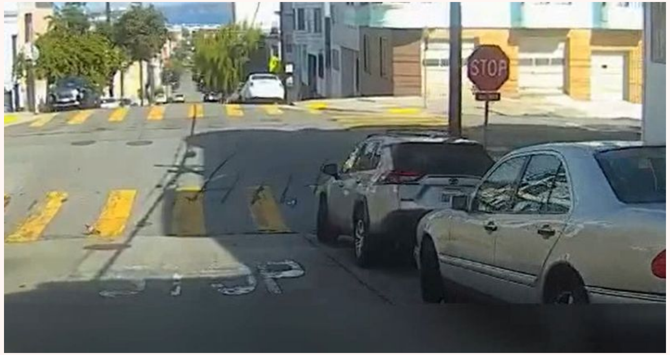
DETAILS

visibility
- ● clear
- ○ partially_occluded
- ○ occluded

Color by: Label | Instance | Group

Objective

To create structured and consistent annotations that capture both the appearance of **stop-related road elements** and the ego vehicle’s behavior in response to **stop signs**



Amit Tal, June 2025

Taxonomy
Labels & Attributes

Label	Type	Attributes	Description (in short)
stop_sign	Bounding Box	visibility	Mark the physical stop sign + define visibility
stop_road_mark	Bounding Box	visibility	Mark the painted “STOP” on road + define visibility
stop_road_line	Polyline	visibility	Mark the stop line marking + define visibility
ego_behavior	Classification (tag)	movement	Define ego car’s motion state

Amit Tal, June 2025

Labels & Attributes stop_sign

STOP
ARRÊT
DUR

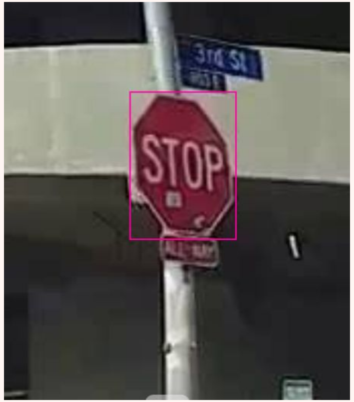
What to annotate: The red octagonal stop sign mounted on a pole

How to draw:

- Use a **rectangle** tightly enclosing the edges of the sign
- Adjust as the angle changes (if rotated or partially visible)
- Continue adjusting the box over the visible part only

Attribute – visibility:

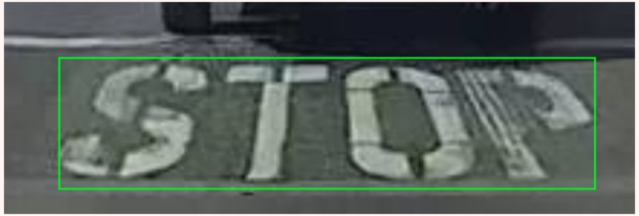
- clear → fully visible with no obstruction
- partially_occluded → partially blocked by trees, poles, vehicles (under 50%)
- occluded → mostly or fully hidden from view (above 50%)



Return to Taxonomy table

Amit Tal, June 2025

Labels & Attributes
stop_road_mark



What to annotate: The painted word “STOP” on the road

How to draw:

- Use a **rectangle** to tightly enclose all letters
- If “STOP” becomes partially visible, continue adjusting the box over the visible part only

Attribute – visibility:

- clear → fully visible with no obstruction
- partially_occluded → partially blocked by trees, poles, vehicles (under 50%)
- occluded → mostly or fully hidden from view (above 50%)

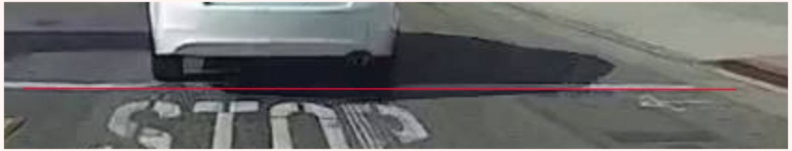
Return to Taxonomy table

Labels & Attributes stop_road_line

What to annotate: The horizontal line where the vehicle is expected to stop.

How to draw:

- Draw a **polyline** from end to end of the visible white stop line
- Place the line as close to the painted line as possible



Attribute – visibility:

- clear → fully visible with no obstruction
- partially_occluded → partially blocked by trees, poles, vehicles (under 50%)
- occluded → mostly or fully hidden from view (above 50%)

Return to Taxonomy table

Labels & Attributes
ego_behavior

What to annotate: The behavior of the ego vehicle (where the dash cam is recording from)

How to classify:

- Use a frame-level tag at any point where movement behavior changes
- No shape drawing is needed

Attribute – movement:

- rolling → moving at a relatively constant pace
- slowing_down → visual deceleration (slower motion of the vehicle is detected)
- stopping → full stop of vehicle (background no longer moves)
- speeding_up → car resumes acceleration

Return to Taxonomy table

Labels & Attributes Additional info (1)

Using Interpolation during annotation work

The annotation platform we use supports **interpolation of tracks**, meaning:

1. You don’t have to manually draw the bounding box/polygon in **every frame**.
2. You place a **keyframe** every few frames, and the interpolation tool will estimate the positions in between

When you can skip frames (use interpolation):

1. Ego vehicle moves forward in a straight line, and the camera is stable
2. The stop sign or road markings remain visible and change position gradually
3. In these cases, label **every 5–10 frames** (depending on object speed and distance)

Labels & Attributes Additional info (2)

When you should NOT skip frames:

1. The ego car is turning, causing sudden shifts in sign angle or field of view
2. The stop sign or markings move quickly **out of frame**, or angle/scale changes sharply
3. During intervals where the object becomes **occluded or reappears** suddenly (due to a tree, pole, or another vehicle)

In these cases:

1. Add **closer keyframes** (e.g., every 1–2 frames).
2. Manually adjust bounding boxes to preserve tightness and accuracy.

Annotation Workflow Summary*

1. **Start frame-by-frame annotation** once the stop sign is visible.

2. **Draw annotations**:
    - Tight rectangle over **stop_sign**
    - Tight rectangle over **stop_road_mark**
    - Polyline for **stop_road_line**

3. **Set visibility attribute** for each visual element

    - Tight rectangle over **stop_sign**
    - Tight rectangle over **stop_road_mark**
    - Polyline for **stop_road_line**

4. Classify **ego behavior** using **ego_behavior** at frames where motion state changes

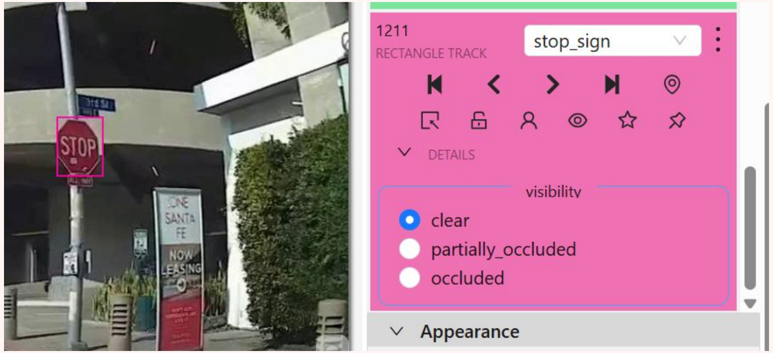
5. Use skipping frames according to requirements in pages 8-9

6. Continue until the ego car has passed the intersection. Once the annotated objects are **not visible** at all in the scene, use ‘end annotation’ to remove it from annotated timeline

*** Assuming annotation work is being done over video clips**

Example*

For task id: 81a9336e-d005-49af-bb75-165180205572
Frame no. 266

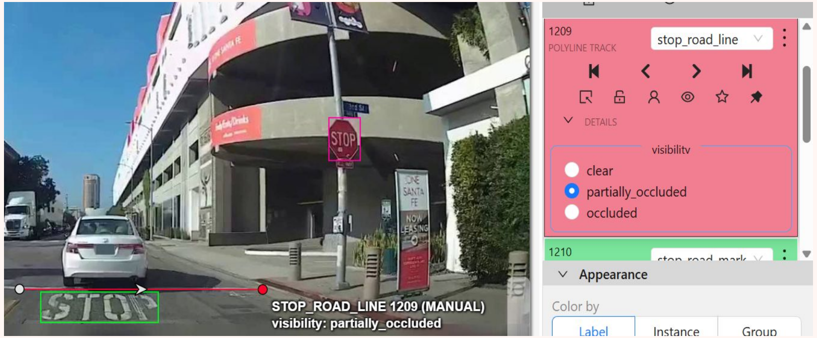


1211
RECTANGLE TRACK
stop_sign

DETAILS

visibility
● clear
○ partially_occluded
○ occluded

Appearance



1209
POLYLINE TRACK
stop_road_line

DETAILS

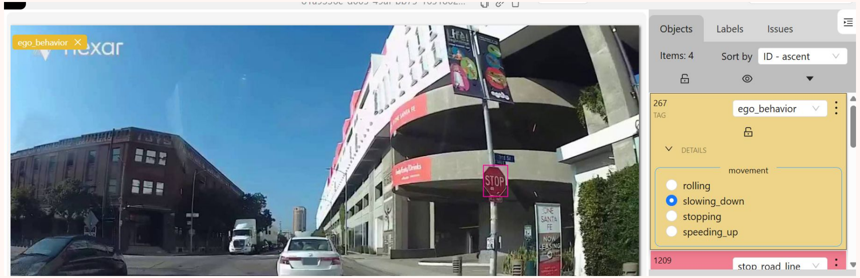
visibility
○ clear
● partially_occluded
○ occluded

1210
stop_road_mark

Appearance

Color by: Label | Instance | Group

STOP_ROAD_LINE 1209 (MANUAL)
visibility: partially_occluded



ego_behavior

Objects | Labels | Issues

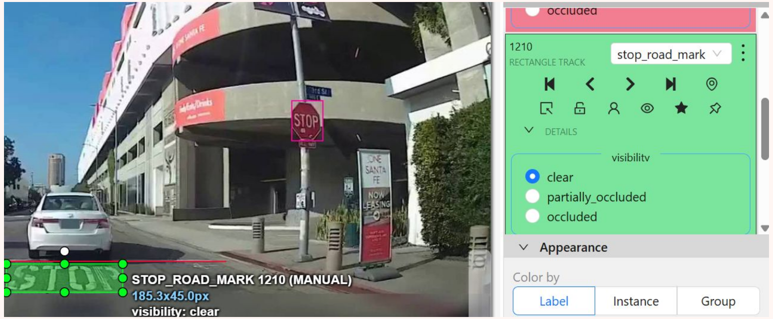
Items: 4
Sort by: ID - ascent

267
TAG
ego_behavior

DETAILS

movement
○ rolling
● slowing_down
○ stopping
○ speeding_up

1209
stop_road_line



occluded

1210
RECTANGLE TRACK
stop_road_mark

DETAILS

visibility
● clear
○ partially_occluded
○ occluded

Appearance

Color by: Label | Instance | Group

STOP_ROAD_MARK 1210 (MANUAL)
185.3x45.0px
visibility: clear

*** Used CVAT for the annotation example + a json file representing the taxonomy used here was submitted to my github branch**

Amit Tal, June 2025

Assumptions

- **Visual-only annotation**: No access to GPS, speed, LiDAR, or inertial sensors. All annotations are based on dashcam footage alone.

- **Ego-vehicle is the camera holder**: The annotator should assume the dashcam is mounted on the ego vehicle; behavior tags are based on the movement of this vehicle.

- **Bounding boxes and polylines reflect object contours**: We're using rectangles for **stop_sign** and **stop_road_mark**, and polylines for **stop_road_line**, assuming a 2D projection view is sufficient.

- **Behavior is inferred from visual motion**: Attributes like **rolling**, **stopping**, and **slowing_down** are based on rough estimations (using the rapid changes of the background from frame to frame).

- **One ego_behavior tag per segment**: The movement label is used once from start to end of the labeling task (segment). The changes made come from the related attribute, and they are frame dependent.

Tradeoffs

- **Skipping frames using interpolation:** Annotators are encouraged to skip frames when the scene is stable and the vehicle moves straight, relying on the interpolation feature. This reduces annotation time (and as a result, cost) without significantly impacting bounding box accuracy. When sharp turns or occlusions occur, annotators return to frame-by-frame work.

- **Visual estimation instead of sensor data:** Due to lack of speed and motion sensors, annotators must infer movement states visually. While this introduces some **subjectivity**, it reflects the kind of perception our models should rely on (as humans do). Nevertheless, by using a limited visibility scale (for consistency purposes), we reduce subjectivity and keep the annotation work objective enough for model training, while acknowledging real-world variability.

- **Focused scope (stop behavior only):** We intentionally exclude labeling **pedestrians, cross traffic, or full vehicle detection** in this phase to maintain focus. This reduces annotation workload while allowing the model to learn the stop sign behavior pattern robustly before expanding to more complex scenarios.

- **No timing or distance labels:** While we do not directly measure stopping time or distance to the stop line, the combination of visibility annotations and ego behavior tags offers a strong proxy signal for learning and predicting stop-related driving behavior.

Amit Tal, June 2025

How This Labeling Scheme Supports Model Training

- **Behavioral signal from visual cues:** Our annotations reflect how a perception-based model "sees" the world - using only visual inputs to infer behavior like stopping, rolling, or accelerating near stop signs.
- **Helps the model learn driver behavior:** By tagging how the car moves (slowing, stopping, rolling), the model can learn what drivers usually do near stop signs.
- **High-quality, low-noise labels:** Narrowing the task to stop-related context and ego behavior allows for cleaner, more consistent labels - essential for generalization and reducing overfitting in early-stage model training.
- **Temporal learning enabled:** Ego behavior tags add a sequence-based dimension to the dataset, enabling training of models that predict not only what is visible, but what the driver will do next.
- **Easy to expand later:** Current setup gives us a strong base. We can always add more later (like pedestrians or other cars) once the model learns the basics well.