Arthur Khozin

Stop sign annotation guide – Nexar

Agenda

- Task description.

- Examples

- Guidelines

- Assumptions

- Labelling scheme justifications.

Task

- In the process of creating an AI model that will analyze the driver behaviour around stop signs we need to create a ground truth data.
- The data is taken as video from dashcams.
- We are going to **place hexagonal tag on the stop sign**.
- Each tag will receive a **label indicating the scenery status**:
  - **Clear** - in cases that stop sign in entirely clear in appearance.
  - **Covered** - in cases there is an object between the camera and the stop sign.
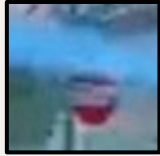
Examples

- There are several types of appearance of stop signs in the videos.
    - **Full appearance.**
        - The entire stop sign is visible.
    - **Partial/Covered appearance.**
        - There is an object on the windshield/ object between the camera and the stop sign.
    - **Taken from distance.**
        - In all the videos we are approaching the stop sign, this allow us to view it at different distances and resolutions.
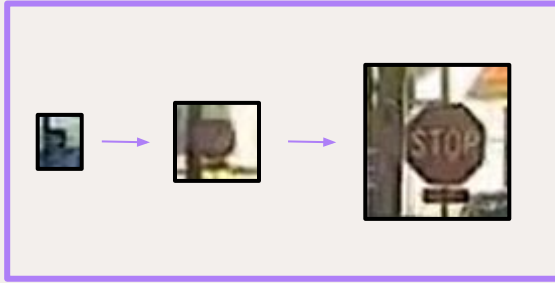
Examples – Full

Examples - Partial

Examples – Taken from Distance

Annotation guidelines

- We are going to go over the video frame by frame and place a bounding box over the stop sign in each frame.
- The bounding box will be in the shape of the stop sign (Hexagon)
- The bounding box will be placed tight within the borders of the the stop sign (examples attached).
- The stop signs can appear from multiple angles and not directly face the camera and should be tagged according to the angle of appearance.
- The stop sign can be covered by objects or interference on windshield, it should be tagged as usual yet place a label "**Covered**"

Label: covered   Label: clear

Label: clear

Assumptions

- I assume we need to tag still images or frames, which can take a lot of time and man hours to complete the task.
    - Tradeoff 1: Taking each 3rd/7th frame per second. Creating a large dataset, while saving man hours.
    - Tradeoff 2: Using a video tagging tool. That can extrapolate to the next frames where the bounding box should be. Yet this can introduce some "noise" in the form of non pure stop sign pixels.
- I assume we will be placing hexagonal tags, while most CV model think in rectangle bounding box, yet this will allow us to train on pure pixels or if needed change the bounding box based on its centroid to rectangle one if needed.
- I assume that for each tag we will need to add a label to indicate the status of the stop sign. Is it full visible? Is it covered by something?, this will allow us to be more robust when training for the road conditions.

Labelling scheme justifications

- My scheme covers the range of appearance of the signs, video is taken from distance and we are getting closer and closer to the sign. This allow us to observe the sign from distance and create tags that reflect different resolutions. Additionally we will observe the signs under interference such as something covering the sign, this is a real world issue and we need to account for this when training the model. In this scheme i introduce a label for each tag to indicate if its "clear" or "covered". Additionally we will place a hexagonal tag on the stop sign to be as tight as possible so we will extract only stop sign features for the model and not features from the surrounding. All of those will allow account for real world issue that our model should be adapted to.