Erchi Zhang

Annotation guideline - Fixplainer

After you have run gen_box.py for your own video, a Json file named "boxes.json" will appear in your designated output path, along with a directory named "img" that contains all the video frames extracted from your uploaded video. In boxes.json, there are many video frame numbers, and inside each video frame numbers, there are many "boxes" denoting the objects detected, with class from 0 to n, indicating different types of objects, including human and dogs. Each of these boxes comes with their coordinates and RGB information, and so forth.

- If you find the object inside the box is correctly tracked, change the class of this box to "0".
- If you find the object inside the box is wrongly tracked (e.g., there is an ID-switch for this object), change the class of this box to "1".
- If you find some missing objects that appears on the video and they should be tracked while they are not appearing as boxes inside this Json file, add their boxes manually, mark them as class "1" and record their coordinates (topx, topy, botx, boty) via https://www.makesense.ai/. You do not need to include any other information for those missing objects other than coordinates.

As stated in our paper section 4.2, when an ID-switch happens, we only count the first frame that the ID-switch happens for this box as a failed tracking. For example, if we have three frames 1, 2, and 3, we only make the class of the object that has this ID-switch in frame 1 as class "1", while keeping its class in frame 2 & 3 and class "0".

Still confusing about how to really annotate the generated Json file?
Don't worry, please refer to our directories that contain many of the annotated data from us. See how we annotated each Json file and how they correspond to each video frame, and you will soon learn how to annotate your own data!