

Evaluation guide

This guide explains how our evaluation should be set up.

Training/validation videos

First we need videos to perform the validation on. Since we will be evaluating training and validation data separately for each category 1-4, we will need 8 videos.

These videos can be found in the *training* and *testing* directories. These videos should follow specific naming conventions depending on the category and whether it is training or testing data. For example the training video for category 1 can be found at *training/training_vid_cat1.mp4* while the testing video can be found at *testing/test_vid_cat1.mp4*.

These 8 videos are created manually by gathering clips from the *dataset/TrainingSet* directory for each category and putting some in the training video and putting others in the validation video. Note that not all videos from the directory will be used, otherwise we would have to label thousands of images.

Getting labels

For evaluating the performance of our localization algorithm, we also need labels for both training and validation data. We store these labels in a json format. To create this json file we have created a labelling tool to facilitate the process. The input of this tool is a video and the output is the json file containing the labels. We run this tool for each of the 8 videos to get 8 label files.

To run this tool you open up a terminal and execute the following command:

```
python label.py --input file-path1 --output file-path2
```

This will take the file at *file-path1* as an input video and output the json file of the labels for that video at *file-path2.json*. If *--output* is not set, it will be defaulted to *labels.json*.

An example of executing this command for training category 1 is:

```
python label.py --input training/training_vid_cat1.mp4 --output training/training_labels_cat1
```

Note that the labels also follow naming conventions. It is the same as the video but “vid” is replaced with “labels”.

Now that you have run the command a window with an image should pop up. You can draw the bounding boxes for the label by holding down the left mouse button. Once you let go the box is drawn. You can draw multiple boxes. If you want to confirm the boxes you have drawn in this frame or want to continue to the next frame, you can press any key. If you are not happy with the boxes you can press right click to reset them for the current frame. Sometimes the boxes will look a bit odd with sides missing but that doesn't impact the labelling, it's just a visual error. Once you have passed over each frame the output json file will be created at the specified location.

Running the evaluation

Running the evaluation is done using a simple command:

```
python localization_eval.py
```

This command also has an optional parameter `--sample_freq some_int` which is the sampling frequency used for evaluation. The default is set to 5 which means that every fifth frame will be used for evaluation. The higher this number is, the faster the evaluation runs, but it comes at the cost of less accurate results.