Supplementary Material - S^{3}AD: Semi-supervised Small Apple Detection in Orchard Environments

Robert Johanson*, Christian Wilms*, Ole Johannsen, and Simone Frintrop
Computer Vision Group, University of Hamburg, Germany
* denotes equal contribution
{firstname}.{lastname}@uni-hamburg.de

This supplementary material includes details on the automated apple property annotation in our dataset MAD (see Sec. 1), details about the training of our TreeAttention module in Sec. 2, and a description of the tile selection process for the semi-supervised training of S^{3}AD's detector (see Sec. 3).

1. Apple Property Annotations in MAD

As briefly mentioned in the main paper, we automatically annotate the apple instances with three properties. Details about the annotation process regarding the properties as well as statistics on the properties are presented below.

1.1. Relative Size

The relative size of an annotated apple instance is directly calculated based on the annotated bounding box. It is defined as the area of the bounding box over the area of the image, which is 3840×2160 for all images in MAD. Hence, the relative size is a value between 0 and 1. Measuring relative size in contrast to absolute size makes this property comparable to other datasets, since images are usually resized at the input stage of object detection systems.

Figure 1a shows the distribution of relative sizes in the test split of our dataset MAD. It is clearly visible that most apples have a relative size below 0.001, resulting in an absolute area below 91^2 pixels in MAD's high-resolution images. Above this level, only a few annotations (8.4%) exist, which indicates a strong focus of MAD on small apples.

1.2. Occlusion

The level of occlusion is measured as the portion of an annotated bounding box depicting the actual apple. To estimate this area, we use the domain knowledge that the apples in our dataset are primarily red, while the occluders (mostly leaves and branches) are green or brown. Therefore, we estimate the hue of the apples by taking the peak hue value (HSV color space) across all pixels of annotated apples in MAD. Subsequently, we apply binary segmentation within each annotated box using the peak hue with some margin. All pixels within each box that are close to the determined peak hue value comprise the actual apple, while the remaining pixels show background areas or occluders. An example of this segmentation is visible in Fig. 2, where green pixels denote estimated occluders within each annotated bounding box. Since we measure the portion of an annotated bounding box depicting the actual apple, the range for the level of occlusion is $0, \ldots, 1$. Here, 0 represents full occlusion, while 1 represents no occlusion.

The distribution of the occlusion levels is depicted in Fig. 1b. From the figure, it is visible that only a few annotated apples are almost occlusion-free (occlusion-level > 0.9), which is also related to the automatic annotation process that counts the corner areas of the bounding box that do not cover apple areas as occluders. Apart from this effect, most apples (84.2%) are moderately occluded, with an occlusion level between 0.3 and 0.9. Severe occlusion with less than a third of the apple being visible only occurs on 10.9% of the annotated apples. This is also related to the difficulty these apples pose for manual annotation.

1.3. Lighting Conditions

Similar to the relative size, we directly derive the lighting conditions from the annotated bounding box. We define the lighting condition as the mean intensity level of the annotated bounding box in the HSV color space. Hence, the lighting condition measures if an annotated apple is dark and in the shadow of a leaf, or directly in the sun and much brighter. Similar to the previous measures, the range for the annotated lighting condition is $0, \ldots, 1$.

Figure 1c shows the distribution of the intensity values as a surrogate for the lighting conditions across MAD's test split. While most apples are well illuminated, there is a general tendency of the distribution towards darker intensities. This indicates that a relevant amount of apples is in the shadow of leaves or other parts of the tree. The fact that almost no annotated apples are very bright or dark is not surprising. Very dark apples are also very difficult to spot