“Cucumber Intelligence” Project Description:

Until spring this year, I had worked on this with my former employer, TITUS GmbH, as part of the “FoodChain” research project (cf. [1]). One part of the project was the surveillance of cucumber harvesting, with the aim of recording instantly how many cucumbers (pickles/gherkins) were harvested.
Generally, this is done by hand, by employees laying flat on a so-called “Gurkenflieger” [2]:

The harvesters put the cucumbers on conveyor belts which transport the cucumbers to the trailer.
In our project, we took photographs (around O(1000), I think) and videos of cucumbers on the belt, and student interns annotated the cucumbers on a subset of these photos.
(Unfortunately, I cannot access our data and results at this moment, because they were lost to the Cloud when our company went bankrupt. Most probably, I will be able to recover the annotated data.)

A colleague of mine developed an AI model (YOLO-based, somehow) that was able to discriminate the cucumbers on the belt. Unfortunately, I only have a picture of the result from an untrained model:

In the end, the cucumber detection results were much better. However, some challenges remained:

1) Since the conveyor belt moves, it is necessary to somehow identify and label individual cucumbers, in order to avoid double counting in time-adjacent photographs. This applies both to moving cucumbers and to cucumbers that may get stuck at the edges of the belt.

2) Cucumbers lying on top of, and thus partially obscuring, each other present another problem for counting.

3) It would be nice to measure the sizes of the harvested cucumbers. Of course, this is difficult now, as we didn't yet include a scale in the photos (I think). But it would be possible to compute a relative size, at least.

4) Now that I think of it, we didn't work on the continuity at all. That is, somehow try to count the number of transported cucumbers, or their integrated volume, over a given amount of time. This should be done!

5) Finally, we only analysed photographs with the AI model. It could be interesting to use video footage instead. Picture quality may be an issue here. I'm not sure about the videos we've taken, and I don't think the cucumbers in the videos have been annotated.

So, my portfolio project idea would be to work on these five challenges, under the assumption that I can retrieve the original data set (labelled and unlabelled). The result might be used by the other institutes/organizations who are still working in the FoodChain project, which is funded at least until the end of this year.

Björn
(bjoern.opitz@posteo.de)

[1] https://www.maz-online.de/lokales/dahme-spreewald/dahme-spreewalds-landwirtschaft-wird-digital-startschuss-fuer-projekt-foodchain-YNK6GA4X4EVOCOV3H47BQQVFLA.html

[2] https://de.wikipedia.org/wiki/Gurkenflieger