Optimizing existing pre-trained convolutional neural networks to predict strawberry maturity levels

By Jurriaan Buitenweg (jbuitenweg@tudelft.nl)

Supervisors: Junhan Wen (junhan.wen@tudelft.nl), Thomas Abeel(t.abeel@tudelft.nl)

Introduction

Background

- Reduce food waste
- Maturity level prediction on a 1-10 scale.
  - 7 or 8 is marketable

The research questions are:

Can convolutional neural networks be used for classifying strawberry maturity levels?

1. Can existing pre-trained convolutional neural networks be used to classify the maturity level on a 1-10 scale given a segment? If so, how accurate?
2. Can existing pre-trained convolutional neural networks be used as a backbone within FasterRCNN to detect strawberries and classify their maturity level directly? If so, how accurate?

Methods

Two experiments will be done:

1. The CNN's are directly applied as a classifier on strawberry segments.
   - Mean squared error
   - Top1 accuracy
   - Top 3 accuracy
   - Cross entropy and early stopping
2. The CNN's will be used as a backbone within the FasterRCNN architecture to find out how they would behave together with object detection.
   - Mean squared error
   - Intersection over union(threshold = 0.6)

- example on IoU in figure 8 [1]

The CNN's used:

- VGG19
- Alexnet
- InceptionV2
- EfficientNetB2
- Resnet50

These will be tested with both stochastic gradient descent(SGD) and adaptive moment estimation (ADAM) as optimizer

Materials

- Two sample sets
  - For experiment 1: Segment with ripeness annotation by three experts on a 1-10 scale (figure 1 and 2)
  - For experiment 2: Full RGB images with bounding box annotation and ripeness annotation by three experts on a 1-10 scale (figure 3)
- Training set: Images of segments/strawberries from August
- Testing set: Images of segments/strawberries from May and June

Figure 1: Strawberry segment

Figure 2: Distorted strawberry segment

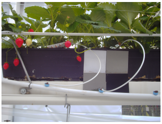

Figure 3: Sample for the faster rcnn experiment

Results

Three best results for experiment 1 sorted on top1 accuracy (figure 4)

- Top1: What percentage is the prediction the exact same as the average of the 3 ratings given by the experts (rounded to nearest integer)
- Confusion matrix indicates off-by-one-errors (figure 5)
- Loss plot shows training process (figure 6)

Model	Top1	MSE	Top3
AlexNet(SGD)	35.20%	1.56	69.00%
AlexNet(ADAM)	32.40%	2.05	63.38%
VGG19(SGD)	30.90%	3.43	59.15%

Figure 4: Comparison on the three best performing CNN's for segment classification.

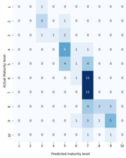

Figure 5: Confusion matrix for AlexNet(SGD)

Three best results for experiment 2 sorted on average intersection over union(IoU) (figure 7)

Model	MSE	Average IoU
ResNet50(ADAM)	2.49	0.73
EfficientNetB2(ADAM)	1.13	0.67
AlexNet(ADAM)	0.00	0.63

Figure 7: Comparison on the three best performing CNN's as a backbone within FasterRCNN.

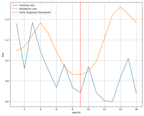

Figure 6: Loss function for AlexNet(SGD) plotted over 16 epochs with the early stopping checkpoint at 9 epochs.

Discussion and conclusion

- Experiment 1
  - Low top1 accuracy
    - Main cause are off-by-one errors
  - Good MSE and high top3 accuracy
    - Model can be used as indication
- Experiment 2
  - Low MSE and good IoU
    - Good classification but detection can be better.
    - Further research needs to be done
- Future research
  - More extensive data collection process
  - More models with different qualities
  - Alternatives to early stopping
  - Different color spaces

[1]https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/

TUDelft