TU WIEN Informatics

Flower State Classification for Watering System

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Software Engineering & Internet Computing

eingereicht von

Tobias Eidelpes, BSc
Matrikelnummer 01527193

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Ao.Univ.-Prof. Dr. Horst Eidenberger

Wien, 20. Februar 2023

Tobias Eidelpes   Horst Eidenberger

TU WIEN Informatics

Flower State Classification for Watering System

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieur

in

Software Engineering & Internet Computing

by

Tobias Eidelpes, BSc
Registration Number 01527193

to the Faculty of Informatics

at the TU Wien

Advisor: Ao.Univ.-Prof. Dr. Horst Eidenberger

Vienna, 20th February, 2023

Tobias Eidelpes — Horst Eidenberger

Erklärung zur Verfassung der Arbeit

Tobias Eidelpes, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 20. Februar 2023

Tobias Eidelpes

Danksagung

Ihr Text hier.

Acknowledgements

Enter your text here.

Kurzfassung

Ihr Text hier.

Abstract

Enter your text here.

Contents

CHAPTER 1

Evaluation

The following sections contain a detailed evaluation of the model in various scenarios. First, we present metrics from the training phases of the constituent models. Second, we employ methods from the field of Explainable Artificial Intelligence (XAI) such as Local Interpretable Model Agnostic Explanation (LIME) and Gradient-weighted Class Activation Mapping (Grad-CAM) to get a better understanding of the models' abstractions. Finally, we turn to the models' aggregate performance on the test set and discuss whether the initial goals set by the problem description have been met or not.

1.1 Object Detection

The object detection model was trained for 300 epochs and the weights from the best-performing epoch were saved. The model's fitness for each epoch is calculated as the weighted average of mAP@0.5 and mAP@0.5:0.95:

$$f_{epoch} = 0.1 \cdot \mathsf{mAP@}0.5 + 0.9 \cdot \mathsf{mAP@}0.5{:}0.95 \tag{1.1}$$

Figure 1.1 shows the model's fitness over the training period of 300 epochs. The gray vertical line indicates the maximum fitness of 0.61 at epoch 133. The weights of that epoch were frozen to be the final model parameters. Since the fitness metric assigns the mAP at the higher range the overwhelming weight, the mAP@0.5 starts to decrease after epoch 30, but the mAP@0.5:0.95 picks up the slack until the maximum fitness at epoch 133. This is an indication that the model achieves good performance early on and continues to gain higher confidence values until performance deteriorates due to overfitting.

Overall precision and recall per epoch are shown in figure 1.2. The values indicate that neither precision nor recall change materially during training. In fact, precision starts to

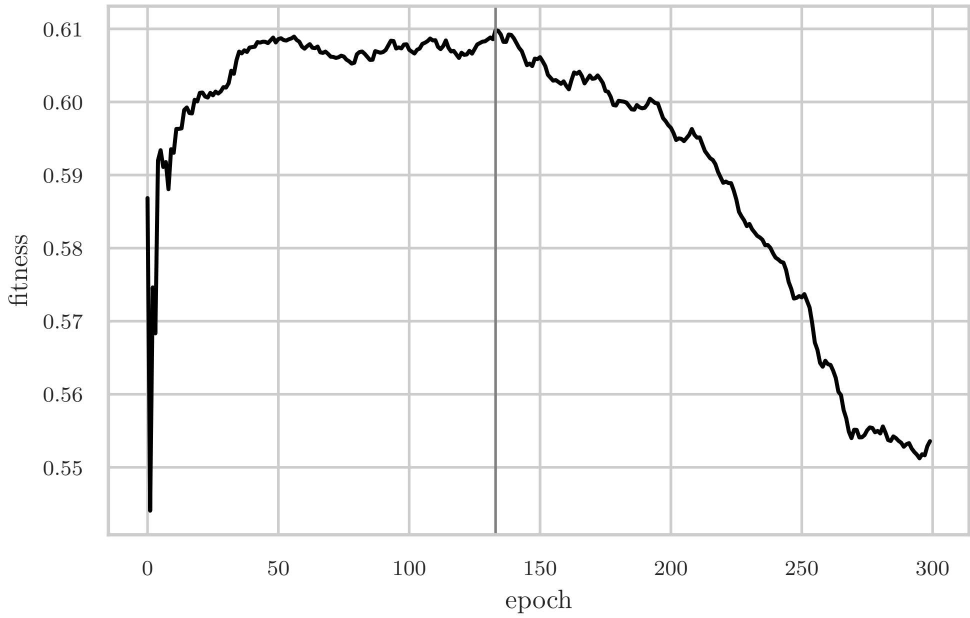

Figure 1.1: Model fitness for each epoch calculated as in equation 1.1.

decrease from the beginning, while recall experiences a barely noticeable increase. Taken together with the box and object loss from figure 1.3, we speculate that the pre-trained model already generalizes well to plant detection. Any further training solely impacts the confidence of detection, but does not lead to higher detection rates. This conclusion is supported by the increasing mAP@0.5:0.95.

Further culprits for the flat precision and recall values may be found in bad ground truth data. The labels from the Open Images Dataset [KRA+] are sometimes not fine-grained enough. Images which contain multiple individual—often overlapping—plants are labeled with one large bounding box instead of multiple smaller ones. The model recognizes the individual plants and returns tighter bounding boxes even if that is not what is specified in the ground truth. Therefore, it is prudent to limit the training phase to relatively few epochs in order to not penalize the more accurate detections of the model. The smaller bounding boxes make more sense considering the fact that the cutout is passed to the classifier in a later stage. Smaller bounding boxes help the classifier to only focus on one plant at a time and to not get distracted by multiple plants in potentially different stages of wilting.

The box loss decreases slightly during training which indicates that the bounding boxes become tighter around objects of interest. With increasing training time, however, the object loss increases, indicating that less and less plants are present in the predicted bounding boxes. It is likely that overfitting is a cause for the increasing object loss from epoch 40 onward. Since the best weights as measured by fitness are found at epoch 133 and the object loss accelerates from that point, epoch 133 is probably the right cutoff

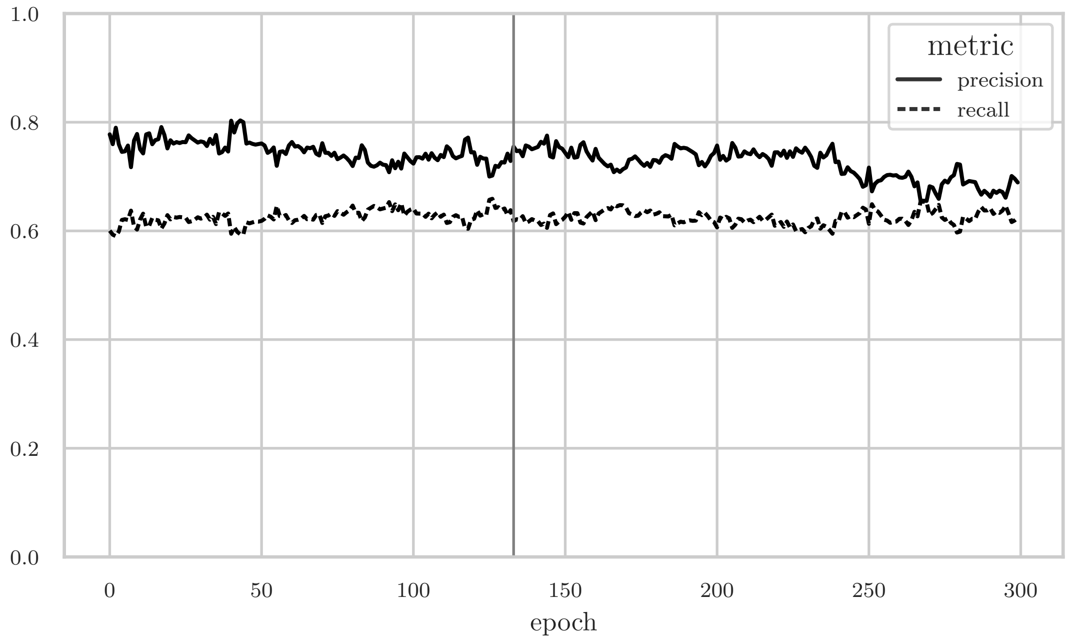

Figure 1.2: Overall precision and recall during training for each epoch.

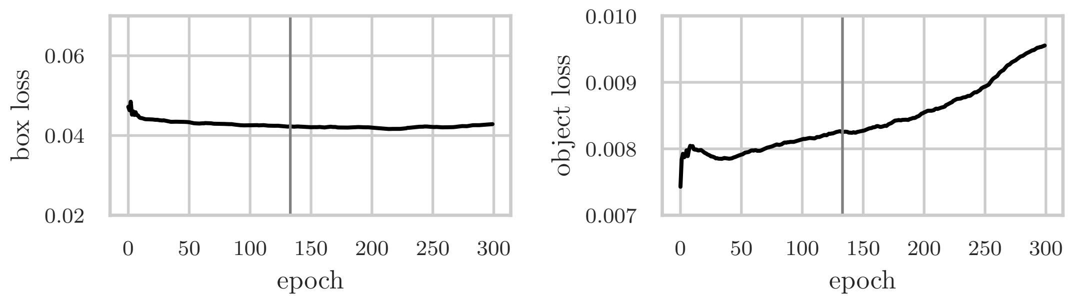

Figure 1.3: Box and object loss[1] measured against the validation set.

before overfitting occurs.

[1]The class loss is omitted because there is only one class in the dataset and the loss is therefore always 0.

List of Figures

List of Tables

List of Algorithms

Bibliography

[KRA+] Alina Kuznetsova, Hassan Rom, Neil Alldrin, Jasper Uijlings, Ivan Krasin, Jordi Pont-Tuset, Shahab Kamali, Stefan Popov, Matteo Malloci, Alexander Kolesnikov, Tom Duerig, and Vittorio Ferrari. The Open Images Dataset V4: Unified image classification, object detection, and visual relationship detection at scale. 128(7):1956–1981.