Paper title: Easy domain adaptation method for filling the species gap in deep learning-based fruit detection

link: https://academic.oup.com/hr/article/doi/10.1038/s41438-021-00553-8/6446655https://www.mdpi.com/40110

Year: 2021

Limitations:

Improve:

Summary:

I. Introduction: There is a vital need in the horticulture research field to understand fruit-related phenotypic traits, and the demand for visual detection techniques in agriculture has increased. Deep learning-based object detection techniques have gradually replaced traditional detection methods and are widely applied in orchard fruit detection. Most related works use a strongly supervised labeling method that requires drawing bounding boxes around the target objects with location and category information for model training. Some works tried to train detection models based on weakly-supervised labeling methods to reduce the labeling cost, such as image-level labels and dot labels, but these methods still required a certain amount of manual data labeling work. Researchers have proposed unsupervised learning methods for agriculture, such as green apple detection in infrared and RGB images and fruit segmentation and localization based on color features. However, unsupervised learning methods did not perform as accurate as supervised learning methods. To address the high dataset labeling cost, some researchers suggest that public available datasets can be used to train fruit detection models. However, the trained fruit detection model showed low generalization ability when applied to real applications. We consider to use Generative Adversarial Networks (GAN) to automatically label different fruit image datasets by only using a set of existing labeled fruit images, and then use the generated images to train several locally good models for each domain based on their own data for fruit detection tasks. The proposed method uses the CycleGAN43 network to transfer the labeled fruit dataset to the unlabeled fruit dataset, and then uses a self-learning method of pseudo labels to improve the labeling accuracy.

II. CycleGAN datasets: The apple2orange dataset contains 995 apple images and 1019 orange images to train an image transformation model. The orange2tomato dataset contains orange images from apple2orange dataset and tomato images collected from the Internet.

a. Object detection datasets: (1) A dataset of 664 orange images was collected from an orange orchard in Sichuan Province, China, and then randomly divided into a training set and a test set according to a 7:3 ratio. A dataset of red and green apples was created based on the MineApple dataset37. 82 images were selected as the experimental test set and were cropped to remove the influence of fallen apples on the ground.

b. Target tomato dataset: The dataset is based on the dataset published by Mu et al.16, and consisted of 598 unlabeled tomato images and 150 labeled tomato images. The orange images and apple images were collected outdoors, and the tomato images were collected indoors.

III. Workflow of the proposed method: We assume that the source fruit dataset comprises the images IS1,lS1,(IS2,lS2), ... ,(ISNS,lSNS) and the target fruit dataset comprises the images IT1,IT2, ... , ITNT. The overall steps of the method are as follows.
Step 1: Import fruit images into CycleGAN testing network, then construct fake apple dataset DF with labeling information of the source fruit dataset Ds.
Step 3: Use the detection box of the real target fruit in the image IT as the pseudo-label information, and then use the self-learning method to improve the accuracy of the labels.
A: Image transformation: The generative adversarial network (GAN) is a popular model for image transformation. It improves the performance of the discriminator network. The CycleGAN network consists of two generator networks and two discriminator networks. The generator networks convert images between two image domains in different directions and the discriminator networks determine the authenticity of images.
To address the problem of large differences in features between fruits of different species, this paper implements feature transfer between fruits using CycleGAN network. The generated fake fruit images are created by combining the original labeling information in the dataset Ds. Examples of source fruit image and generated fake fruit image at different shooting distances.
B: Fruit detection network: This study applied an improved Yolov3 model to detect small-scale fruit. The model uses a Feature Pyramid Network(FPN) network structure to fuse the deep and shallow network features.
C: Pseudo-label generation: This paper proposes a pseudo-labeling approach to generate labels in the dataset DTu automatically. The model M2 can detect real target fruits and can be used to generate the labeling information in the dataset DTL automatically.
D: Pseudo-label self-learning: The detection bounding box obtained by the model M2 in real target fruit images IT is used as a pseudo label, and the impact of noise in pseudo labels is reduced. This paper proposes a pseudo-label self-learning method that includes a pseudo-label noise filtering operation and a cyclic update operation to reduce the effects of pseudo-label noise, thereby improving the labeling accuracy of pseudo labels. The pseudo label of the real target fruit dataset is calculated by calculating the average score of all detection boxes and filtering out the detection boxes below the average score. Pseudo-label cycle update: The method in this study re-obtains the detection box of the unlabeled real target fruit dataset DTu by using the current fruit detection model M2 at certain intervals of training epochs to improve the labeling accuracy.

Iv. Experimental setup: This experiment uses a computer platform with an Intel Core i7-8700K CPU processor, GeForce GTX 1080Ti GPU graphics card, and ubuntu18.04LTS to implement deep learning models. CycleGAN model was trained using a mini-batch adaptive moment estimation (Adam) optimizer with a momentum factor of 0.5 and a batch size of one. Improved-Yolov3 model training: A stochastic gradient descent with a mini-batch with a momentum factor was used to train the network, and the learning rate was adjusted using the cosine annealing function.

v. Evaluate metrics: This paper uses Precision, Recall, F1 score, and mAP to evaluate the detection performance of the Improved-Yolov3 model. The standard intersection-over-union threshold value of 0.5 was adopted.
a. Evaluation of datasets DS and DF: In this study, the fruit detection model Improved-Yolov344 was trained and tested using the datasets DS and DF. The model achieved a high detection accuracy of 95.1% in the dataset DS_orange and 94.8% and 96.7% in the dataset DF.
b. Adding pseudo labels obtained through different confidence thresholds: This experiment was conducted to compare the test results of real apple images and real tomato images using a model fine-tuned using a pseudo-labeling method. The results show that the model M2 performs best

under the current optimal confidence threshold parameters. When the confidence threshold is exceeded, the mAP value of the model decreases as the confidence threshold value increases, because the low number of pseudo label with high threshold leads to a decrease in the diversity of features learned, which affects the generalization ability of the model.

c. Pseudo-label self-learning method to reduce noise labels: The effect of noise in the acquired pseudo labels is investigated. It is found that the model mAP value increases as the set confidence threshold increases, and decreases thereupon, mainly due to the effect of the confidence threshold on the quality and quantity of the generated pseudo labels.

vi. Generated datasets labels: The proposed method can generate higher quality label data automatically, and the obtained pseudo-labels can well surround the target apples at different locations in the image, which verifies the effectiveness of the proposed method in this study. The model detects apples and tomatoes in real scenarios and can be used to automatically label the unlabeled fruit dataset.

vii. Discussion: This paper proposed a new solution to overcome the current problem of high labeling cost for training data acquisition: the automatic labeling of unlabeled fruit datasets. Public resources for images of fruit species are currently available, and the method in this paper could be used to automatically label other datasets, thereby saving time and improving fruit inspection efficiency. In order to apply this method in practice, the source fruit and target fruit species should have small differences in shape and size, and the background color features and the fruit color features should be distinguished as clearly as possible.

viii. Conclusion: This paper proposed a method for filling the species gap in deep learning - based fruit detection, which can be applied for the acquisition of labeling information from unlabeled target fruit datasets. We intend to study further on the following aspects: 1) how to solve the image transformation problem; 2) how to obtain the best confidence threshold.