University of BRISTOL

Dissertation Research Report

Machine Vision in Agriculture: Fruit Counting and Yield Estimating Model

Presented by:
Pavan Kumar Kavvuri

Supervisor:
Dr Mark Hansen

MSc Robotics

AENGM0029 : Robotics Research Preparation

March 27, 2020

Contents

List of Figures

List of Tables

Abstract

An accurate and reliable image based fruit detection and counting system is crucial for supporting higher level agriculture tasks such as yield mapping and robotic harvesting. This is important for growers as it facilitates efficient utilisation of resources and improves returns per unit area and time. This project aims at developing algorithms for detecting and counting fruits from digital images gathered by the agriculture team at the Australian Centre for Field Robotics, The University of Sydney, Australia.

Introduction

Robotics and Autonomous Systems (RAS) are set to transform global industries. These technologies will have the greatest impact on large sectors of the economy with relatively low productivity such as Agri-Food [Duckett et al., 2018]. The advent of autonomous system architectures gives us the opportunity to develop a new range of flexible agricultural equipment based on small, smart machines that reduces waste, improves economic viability, reduces environmental impact and increases food sustainability. Sensory data collected by these platforms in the field can further provide a wealth of information about soil, seeds, livestock, crops, flowers, fruits, costs, yield, farm equipment and the use of water and fertiliser to help farmers analyse data on weather, temperature, moisture, prices, etc., and provide insights into how to optimise yield, improve planning, make smarter decisions about the level of resources needed, and determine when and where to distribute those resources in order to minimise waste and increase yields[vie,]. This is termed as Precision Agriculture which concerns the use of monitoring and intervention techniques to improve efficiency, realised in application through the deployment of sensing technologies and automation.

Machine Vision offers significant opportunities enabling such intervention, for example, crop monitoring, classifying when individual plants are ready for harvest[Barnes et al., 2010], operating at nights, quality analysis, detecting the onset of diseases etc. Despite these advances in agricultural automation that increased productivity by reducing manual labour and production costs, few agricultural tasks are still being handled manually, challenging the consistently shrinking and increasingly costlier agricultural labour force[Kapach et al., 2012]. Probably, harvesting is the process that has received the least amount of technological development for satisfactory automation.[Jiménez et al., 2000].

Harvesting of delicate fruits like oranges,mangoes, apples or peaches for the fresh market, is a process that cannot be performed using aggressive methods such as shakers. Currently the number of fruits on trees in a mango/apple orchard is estimated via a manual count of a small number of trees to predict resource requirements for harvest, and to arrange marketing[Payne et al., 2013]. Usually about six weeks prior to harvest when the fruits are in stone-hardening stage, the crop load is estimated. At this stage, fruits may be half green and half pale orange colour, or all green.[Payne et al., 2014]. Mangoes are generally harvested at physiologically mature stage and ripened for optimum quality. The best way to observe maturity in mango is the colour of the pulp, which turns cream to light yellow on maturity and hardening of stone[man,].

Three main problems that hinder the automation here are: firstly, building an autonomous navigating robot. Secondly, detecting and localising the ripen fruits on the tree and thirdly, selectively harvesting without harming other fruits. This report limits its scope to the study of fruit detection and counting to accurately predict the yield and exclusively discusses this problem in the context of mango and apple orchards.

The remainder of this report is organised to clearly define the Aims and Objectives, Motivation, Related work, Risks associated and Timeline for this project.

Aims and Objectives

This project will focus on modern data processing and segmentation techniques to facilitate precision agriculture in orchards. By improving yield and fruit count estimates, growers can plan for harvest and also get information about the health of their crops. In the long run, this gives growers valued feedback about how efficiently they use resources, and how they can minimise the use of pesticides and fertilisers to maximise their harvest[Stein, 2016].

The overall research goal is to design and develop a machine learning algorithm for rapid and accurate apple/mango yield estimation. Such systems reduce labor intensity, and increase work efficiency by applying computer vision-based, fast data acquisition. At this stage of the research, we focus on two specific objectives: (1) develop major algorithm modules for Instance based segmentation for yield estimation; (2) conduct performance tests on the models developed. Questions to be addressed in while targeting the objectives are:

1. How can we label the ripeness of the fruit in data?
2. Will the algorithm disclose an occluded fruit?
3. How can we precisely localise the fruit despite heavy leaf occlusions?
4. How can we estimate the yield by just counting fruits in the images?

Motivation

Crop yield estimation is a crucial task in an orchard management. Accurate yield prediction helps growers improve fruit quality and reduce disbursements by making better decisions on size of the harvest labor. Besides, it benefits the packing industry because managers can use estimation results to optimize packing and storage capacity [Wang et al., 2012]. Current industry practice to estimate orchard fruit yield is by workers manually counting fruits in multiple sampling locations. This process is time-consuming and labor-intensive, and therefore the limited sample size is typically not enough to reflect the yield distribution across the orchard, especially in those with high spatial variability[Wang et al., 2012]. Therefore, the present yield estimation practice is inaccurate and inefficient, and improving it might be a major result to the industry. The advantage of instead using a machine vision based system, is that it can facilitate fruit count estimates from a larger number of trees, could be used at individual trees, and at several times during the crop growth period.

Related Work

Major works presented in the literature address the problem of fruit detection as an image segmentation problem (i.e., fruit vs. background). Due to high variation in the appearance of the fruits in field settings, including colour, shape, size, texture and reflectance properties, developing a fast and reliable fruit detection and segmentation system is challenging. Furthermore, in majority of these settings, the fruits are partially abstracted and subject to continually-changing illumination and shadow conditions [Sa et al., 2016]. Here we discuss the most recent advances in AI for Instance segmentation and counting.

Instance Segmentation and Counting:

Instance segmentation is very important in a variety of applications such as autonomous driving, image captioning, and visual question answering. It differentiates the individual object instances in a scene. Such an approach is ideal for counting the fruits in any given image. [Ren and Zemel, 2017] proposes an end-to-end recurrent neural network (RNN) architecture with an attention mechanism to model a human-like counting process, and produce detailed instance segmentation.

Cell counting is most commonly a manual task and can be time-intensive due to overlapping cells, existence of multiple focal planes, and poor imaging quality. [Hernández et al., 2018] proposes a convolutional neural network approach using feature pyramid networks (FPN) with a VGGstyle neural network, for segmenting and subsequent counting of cells in a given microscopy image. [Van Valen et al., 2016] developed DeepCell, which treats the segmentation task as classification problem on a pixel-by-pixel basis and produces fairly low-resolution segmentation masks.

[Kumar and Domnic, 2019] proposed an efficient method to extract the leaf region and count the number of leaves in digital plant images that is used to record the plant growth, plant yield plant width and tallness, leaf area, etc. This involves the counting number of leaves in the plant image by applying Circular Hough Transform (CHT). [Cholakkal et al., 2019] proposed an image level supervised approach that provides both the global object count and the spatial distribution of object instances by constructing an object category density map.

Pixel-wise classification was studied for identifying individual mangoes by the analysis of local colours and textures followed by the blob extraction in [Payne et al., 2014] and Convolutional Neural Networks(CNNs) apple classification followed by Watershed Segmentation [Sa et al., 2016]. [Wang et al., 2012] examined the issue of apple detection for yield prediction by developing a system that detects apples based on their colour and distinctive specular reflection pattern. Region based Convolutional Neural Networks (R-CNN) [Girshick et al., 2016], which combine the RoI approach with CNNs, have produced state-of-the-art detection results on PASCAL-VOC detection dataset [Everingham et al., 2010]. [Sa et al., 2016] successfully demonstrated the use of Faster R-CNN for sweet pepper and rock melon detection in a greenhouse.

Semantic image segmentation performs this densely, resulting in a pixel-wise classification over the image. Applying post-processing techniques can then differentiate individual whole-objects of interest as groups of adjacent pixels. On the other hand, the detection search space is reduced using low-level image analysis to spot regions of interests (RoIs) within the image (e.g. possible fruit regions), followed by high-level feature extraction and classification [Bargoti and Underwood, 2017].

Other works in image segmentation and detection include Global Mixture of Gaussians (GMOG) algorithm with background modeling in RGB color was applied on the real-time video image sequences of apples captured in a greenhouse which correctly identified 85-96 percent of both red and yellow apples and performed at 14-16 frames per second [Tabb et al., 2006]. The potential advantages of using video processing includes allowing harvesting on-the-go, achieving a faster harvest time. Hyperspectral imaging was used to green apple yield estimation, as it gives wealth of information both in the visible and the near-infrared (NIR) regions. Leveraging several techniques like principle components analysis (PCA) and extraction and classification of homogenous objects (ECHO) for analyzing hyperspectral data, morphological operations, watershed, and blob analysis, a multistage algorithm was developed [Safren et al., 2007]

Risks Register

Table 1: Risk Register

#	Risk	Likelihood	Impact	Mitigation	Score
1	Poor Data quality	2	4	Careful selection	2
2	Incorrect Data Annotation	1	4	Selection of annotated data as pixel-wise manual annotation would be difficult	4
3	Data Insufficiency	3	3	Data augmentation	9

Timeline

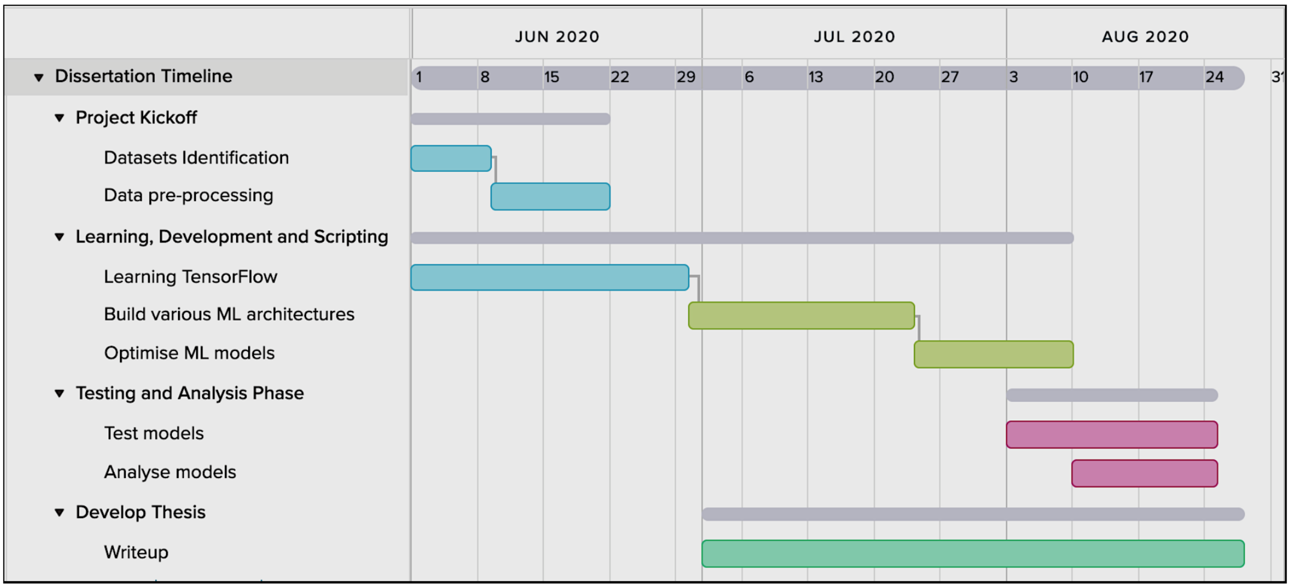

Figure 1: Timeline

References

[man,] Mango Harvesting, howpublished =http://nhb.gov.in/pdf/fruits/mango/man004.pdf/.

[vie,] Views and forecasts about robotics for the ag industry, howpublished =https://robohub.org/views-and-forecasts-about-robotics-for-the-ag-industry/.

[Bargoti and Underwood, 2017] Bargoti, S. and Underwood, J. (2017). Deep fruit detection in orchards. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3626–3633.

[Barnes et al., 2010] Barnes, M., Duckett, T., Cielniak, G., Stroud, G., and Harper, G. (2010). Visual detection of blemishes in potatoes using minimalist boosted classiers. *Journal of Food Engineering*, 98.

[Cholakkal et al., 2019] Cholakkal, H., Sun, G., Khan, F. S., and Shao, L. (2019). Object counting and instance segmentation with image-level supervision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 12397–12405.

[Duckett et al., 2018] Duckett, T., Pearson, S., Blackmore, S., Grieve, B., Wilson, P., Gill, H., Hunter, A., and Georgilas, I. (2018). *Agricultural Robotics: The Future of Robotic Agriculture*. UK-RAS White Papers. UK-RAS Network, UK United Kingdom.

[Everingham et al., 2010] Everingham, M., Van Gool, L., Williams, C. K., Winn, J., and Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88(2):303–338.

[Girshick et al., 2016] Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2016). Region-based convolutional networks for accurate object detection and segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(1):142–158.

[Hernández et al., 2018] Hernández, C. X., Sultan, M. M., and Pande, V. S. (2018). Using deep learning for segmentation and counting within microscopy data. *arXiv preprint arXiv:1802.10548*.

[Jiménez et al., 2000] Jiménez, A., Ceres, R., and Pons, J. (2000). A survey of computer vision methods for locating fruit on trees. *Trans. ASAE*, 43.

[Kapach et al., 2012] Kapach, K., Barnea, E., Mairon, R., Edan, Y., and Ben-Shahar, O. (2012). Computer vision for fruit harvesting robots - state of the art and challenges ahead. *IJCVR*, 3:4–34.

[Kumar and Domnic, 2019] Kumar, J. P. and Domnic, S. (2019). Image based leaf segmentation and counting in rosette plants. *Information Processing in Agriculture*, 6(2):233–246.

[Payne et al., 2013] Payne, A., Walsh, K., Subedi, P., and Jarvis, D. (2013). Estimation of mango crop yield using image analysis – segmentation method. *Computers and Electronics in Agriculture*, 91:57–64.

[Payne et al., 2014] Payne, A., Walsh, K., Subedi, P., and Jarvis, D. (2014). Estimating mango crop yield using image analysis using fruit at 'stone hardening' stage and night time imaging. *Computers and Electronics in Agriculture*, 100:160 – 167.

[Ren and Zemel, 2017] Ren, M. and Zemel, R. S. (2017). End-to-end instance segmentation with recurrent attention. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6656–6664.

[Sa et al., 2016] Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., and McCool, C. (2016). Deepfruits: A fruit detection system using deep neural networks. *Sensors*, 16(8):1222.

[Safren et al., 2007] Safren, O., Alchanatis, V., Ostrovsky, V., and Levi, O. (2007). Detection of green apples in hyperspectral images of apple-tree foliage using machine vision. *Transactions of the ASABE*, 50:2303–2313.

[Stein, 2016] Stein, M. (2016). Improving image based fruitcount estimates using multiple view-points.

[Tabb et al., 2006] Tabb, A., Peterson, D., and Park, J. (2006). Segmentation of apple fruit from video via background modeling.

[Van Valen et al., 2016] Van Valen, D. A., Kudo, T., Lane, K. M., Macklin, D. N., Quach, N. T., DeFelice, M. M., Maayan, I., Tanouchi, Y., Ashley, E. A., and Covert, M. W. (2016). Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. *PLoS computational biology*, 12(11).

[Wang et al., 2012] Wang, Q., Nuske, S., Bergerman, M., and Singh, S. (2012). Automated crop yield estimation for apple orchards. volume 88.