PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Melon yield prediction using small unmanned aerial vehicles

Tiebiao Zhao, Zhongdao Wang, Qi Yang, YangQuan Chen

Melon yield prediction using small unmanned aerial vehicles

Tiebiao Zhao[a], Zhongdao Wang[b], Qi Yang[c], and YangQuan Chen[a]

[a]University of California, Merced, 5200 Lake Rd, Merced, California, USA
[b]Tsinghua University, 30 Shuangqing Rd, Haidian Qu, Beijing, China
[c]Shenyang Ligong University, 6 Nanping Middle Rd, Hunnan Qu, Shenyang, Liaoning, China

ABSTRACT

Thanks to the development of camera technologies, small unmanned aerial systems (sUAS), it is possible to collect aerial images of field with more flexible visit, higher resolution and much lower cost. Furthermore, the performance of objection detection based on deeply trained convolutional neural networks (CNNs) has been improved significantly. In this study, we applied these technologies in the melon production, where high-resolution aerial images were used to count melons in the field and predict the yield. CNN-based object detection framework-Faster R-CNN is applied in the melon classification. Our results showed that sUAS plus CNNs were able to detect melons accurately in the late harvest season.

Keywords: sUAS, Faster R-CNN, melon detection, yield

1. INTRODUCTION

Watermelon, cantaloupe and honeydew are the three predominant melon varieties. The combined value of all melons is among the top three highest of vegetable and melon crop in the United States.[1] Consumption of melon is among top five ranking vegetable and fruit crops in the United States. Figure 1 shows the production and imports of melons in the United States from 1970 to 2012. The imported melons have increased in the past few decades and it has reached around 30 percent of the whole disappearance.[2] Among the imported melons,

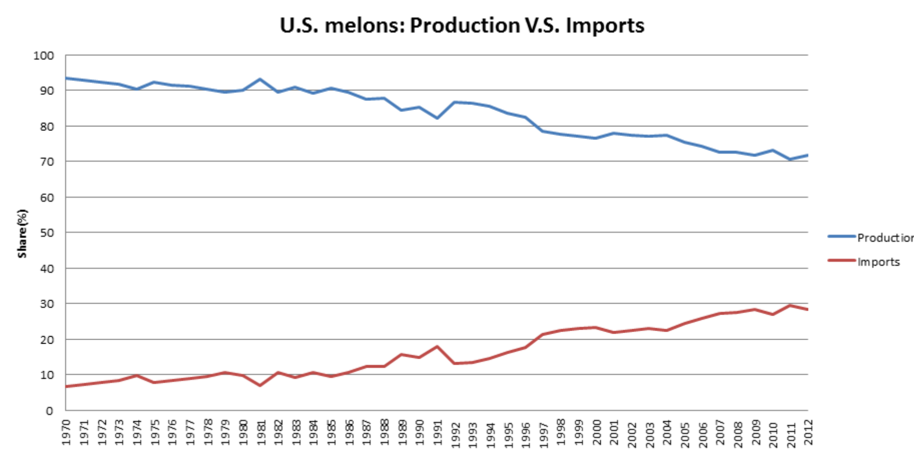


Figure 1. Melons productions and imports in United States from 1970 to 2012.

about 20 percent of import volume is watermelons, which makes the United States the worlds largest importer of watermelons.[3]

Further author information: (Send correspondence to YangQuan Chen)
Tiebiao Zhao: E-mail: tzhao3@ucmerced.edu
Zhongdao Wang: E-mail: wcd13@mails.tsinghua.edu.cn
Qi Yang: E-mail: 9501133@163.com
YangQuan Chen: E-mail: ychen53@ucmerced.edu

Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping II, edited by J. Alex Thomasson, Mac McKee, Robert J. Moorhead, Proc. of SPIE Vol. 10218, 1021808 · © 2017 SPIE · CCC code: 0277-786X/17/$18 · doi: 10.1117/12.2262412

Yiled prediction and monitoring are important for growers of fruits or vegetables, such as melons. The market of melons is subject to lots of factors such as weathers, supplies and prices etc. Yield prediction not only helps growers make better harvest and market plans, better labor management, bu also provides field variability information necessary for future nutrient and water management.

Typically, vegetation indices extracted from remote sensing based images from satellites, manned and unmanned aircrafts are used to monitor yields. Quite a lot studies have been published on this methods. In Ref. 4, NDVI utilizing MODIS (250m) were considered to predict corn yield in four agricultural statistics districts (ASD) in Iowa state between years 2000 to 2014. Aera under the curve of NDVI series showed a significant correlation with yield, with coefficient of deterministic (R-square) over 0.80 in all ASDs. In particular, this model was tested using the data of 2015 and the predictability error was between 5 to 7 percent. In Ref. 5, soil information and high resolution images were used to predict wheat yield. The average performances of three methods counter- propagation artifical neural networks (CP-ANNs), XY-fused networks (XY-Fs) and supervised kohonen networks (SKNs) are 78.3%, 80.92%, and 81.65% respectively. In Ref. 6, grassland biomass estimation models were developed using two spectral bands (red band and near-infrared (NIR)) from satellite images (MODIS). Adaptive neuro-fuzzy inference system (ANFIS) model showed improved estimation compared to multiple linear regression (MLR) and artificial neural network (ANN). In Ref. 7, vegetation health index (VCI) and temperature condition index (TCI) based on advanced very high resolution radiometer (AVHRR) were used to predict potato yield from 1980 to 2014. It was shown that artificial neural network (ANN) model generated the prediction error less than 10%. In Ref. 8, UAV based remote sensing were applied to evaluate the impact of different nitrogen treatments on maize yield. Wide dynamic range vegetation index (WDRVI) NDVI and crop height showed no significant response to extra N application beyond the economic optimum rate.

Figure 2. The orthomosaic of study site, stitched from aerial images taken at the altitude of 60 m above ground

More recently, with the advances of image processing techniques, especially deep learning based algorithms, fruit detection and direct counting become possible. In Ref. 9, deep convolutional neural networks (CNNs) based framework- faster region-based CNN (Faster R-CNN)[10] were used to detect fruits directly in trees with images of colour (RGB) and NIR bands. This model was also retrained and tested on sweet peppers, apples, avocados, mangoes and oranges. To our best knowledge, there are no studies talking about how to predict melon yield by detecting and counting melons directly in the field using UAVs and Faster R-CNN methods.

2. MATERIALS AND METHODS

2.1 Study Site

This research was conducted in a commercial melon orchard of 30 acres in Merced, California, USA ($37°09'47.1"N$, $120°47'30.7"W$). The field were divided in to three blocks, where melons were planted in May, June, July, 2015 respectively. The block we used was grown in the beginning of July, 2015. The first harvest was on September,

(a) Image sample 1 taken at one meter above ground (b) Image sample 2 taken at three meters above ground

Figure 3. Sample1 and Sample2

19th,2015 and the second harvest was a week later, September 26th, 2015. The field was irrigated twice using furrow irrigation.

2.2 Image Collection

Images were collected using the commercial-off-the-shelf (COTS) red-green-blue (RGB) camera ELPH110HS (Canon,Japan). It has a resolution of 4608x3459 pixels. Since our purpose was to detect melons, so only JPG images were taken for further processing. The camera was attached on the bottom of the UAV, built from scratch using Quadkit (3DRobotics, USA). The camera was configured to be triggered by the autopilot on the UAV with the highest speed–one frame/second. The flight altitude was between one and five meters above the ground.

2.3 Faster R-CNN

Great advances have been achieved with the help of region proposal methods[11] and region-based convolutional neural networks (R-CNNs).[12]Fast R-CNN[13] increase the speed by sharing convolutions across proposals. Yet, region proposals are still computationally expensive in object detection systems. Faster R-CNN[14] was proposed to overcome this problem by computing proposals with a deep convolutional neural network–region proposal networks (RPNs). RPNs can predict regions with a wider range of scales and aspect ratios. Most importantly, the features learned in the region proposal step are shared with detection network, reducing the computation cost and enabling near real-time detection.

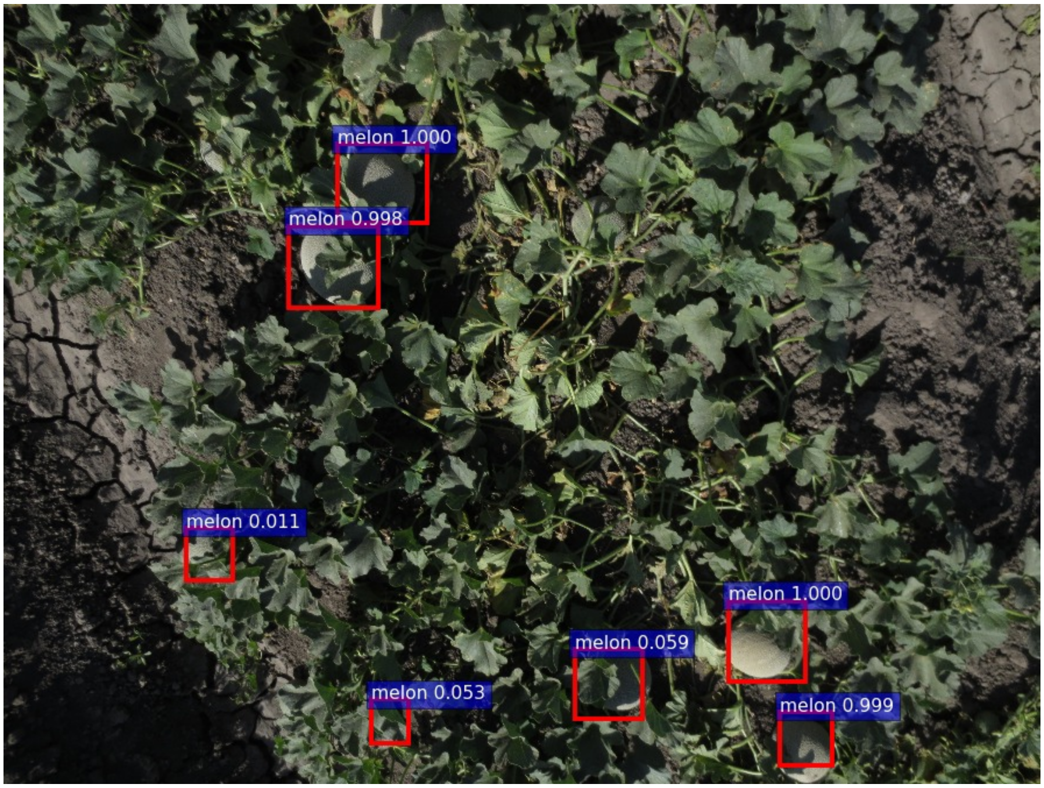
3. RESULTS AND DISCUSSIONS

3.1 Experiment Setup

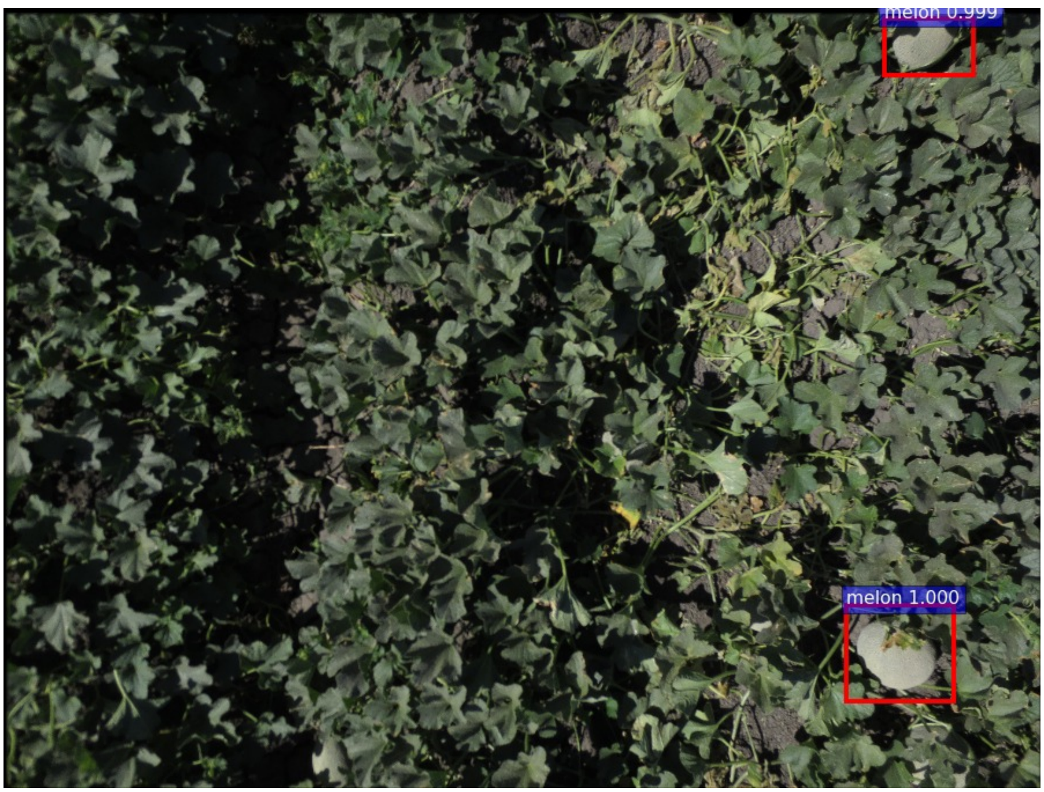
We manually annotated melons in the images with bounding boxes, giving out the left-top corners and right-bottom corners coordinates of the bounding boxes. There are 508 images and 1924 melons in this dataset, about 4 melons per image in average. The dataset is divided into training/validation/test sets, with 350/50/108 images and 1384/144/396 melons in each set. Figure. 3(a) shows the sample image collected at one meter above the ground, and Figure. 3(b) was taken at three meter above the ground.

3.2 Training Details

Compared to the large scale object detection problem, our task is easier because there is only one foreground category to be detected. Therefore we chose the ZF[15] net as our RPN, rather than bigger networks like VGG-16 and VGG-19. Alternating training strategy was adopted for training the RPN and Fast R-CNN, which are the two components of the Faster R-CNN. We first trained RPN, and then used the region proposals to train Fast R-CNN. The network tuned by Fast R-CNN was then used to initialize RPN, and this process was iterated.



(a) Detection test on image 1



(b) Detection test on image 2

Figure 4. Melon detection test using the trained model

In the RPN training stage, the initial learning rate was set to 0.0002 and decayed by one tenth every 60000 iteration. In the Fast R-CNN training stage, the initial learning rate was set to 0.0001 and decayed by one tenth every 60000 iteration. A ZF net model pre-trained on ImageNet dataset was used for initializing the RPN.

3.3 Evaluation

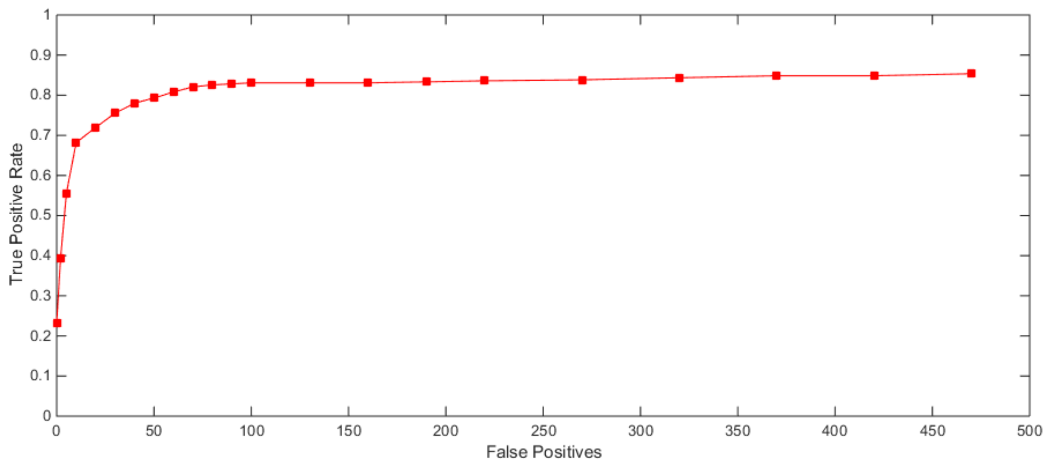


Figure 5. The performance curve of our model in melon detection

To evaluate the detection models performance, we presented a discrete manner for scoring detections in an image. If the ratio of the intersection of a detected region with an annotated melon region is greater than 0.5, a score of 1 is assigned to the detected region, and 0 otherwise. Thus we can get the True Positive Rate

$$TPR = \frac{number\ of\ positive\ samples\ detected}{number\ of\ all\ positive\ samples} \tag{1}$$

and the number of False Positives. Every detected object has a confidence score. By thresholding this score we can get pairs of TPR-FP values. Then the detection performance curve could be plotted with the False Positives as the horizon axis and the True Positive Rate as the vertical axis, as shown in Fig. 5. We can observe that 79.29% of the melon targets are accurately detected when the overall number of False Positives on the whole testing set is 50. Fig. 4 shows the detection performance of our trained model, where red rectangles are bounding boxes for the detected melons and blue texts are the detection scores.

4. CONCLUSIONS

We discussed melon detection in the field using UAVs under different conditions, such as different altitude and different cover rate by leaves or shade. Faster R-CNN performed up to 79.29% recall when the overall number of False Positives was 50 with the collected field dataset, while keeping high detection speed and low requirement of ground truth annotation. This is more competitive than the best accuracy 60% (TPR) we obtained using traditional methods combining features of texture and color.[16]

In developing this model, we selected ZF net instead of VGG-16 or VGG-19, in particular for melon detection in the field, because it reduces the size of networks and hence the possibility of over-fitting and saves the training time. We performed fine-tuning of ZF network based on the pre-trained ImageNet model.

Future work includes continuous effort to collect images so the model could be trained better to handle more complex conditions. At the same time, this paper just discusses melon detection within one image. To extend this work to field level, same melons appearing in the successive frames need to be detected to avoid multiple counts. It is also necessary to determine the size of melons to better evaluate the yield, which is not supported by Faster R-CNN based methods. Furthermore, fractional calculus, as a more general tool, might help improve the performance of neural networks, as shown by its great success in image processing.[17]

ACKNOWLEDGMENTS

Thanks go to Big Idea, Berkeley for funding the project "SmartMelonDrone". Thanks go to Scott Stoddard and Legend Produce for helping coordinate the test filed. Thanks go to Yoni Shchemelinin, Andreas Anderson, Jacob Clark for helping collect flight images.

REFERENCES

[1] Agricultural Marketing Resource Center, "A national information resource for value-added agriculture." http://www.agmrc.org/commodities-products/vegetables/melons/. (Accessed: March 10th,2017).

[2] United States Department of Agriculture Economic Research Service, "The Fruit and Tree Nuts Yearbook." https://www.ers.usda.gov/data-products/fruit-and-tree-nut-data/yearbook-tables/. (Accessed: March 10th,2017).

[3] United States Department of Agriculture Economic Research Service, "The Fruit and Tree Nuts Trade." https://www.ers.usda.gov/topics/crops/fruit-tree-nuts/trade.aspx/. (Accessed: March 10th,2017).

[4] Shrestha, R., Di, L., Eugene, G. Y., Kang, L., Li, L., Rahman, M. S., Deng, M., and Yang, Z., "Regression based corn yield assessment using modis based daily ndvi in iowa state," in [*Agro-Geoinformatics (Agro-Geoinformatics), 2016 Fifth International Conference on*], 1–5, IEEE (2016).

[5] Pantazi, X. E., Moshou, D., Alexandridis, T., Whetton, R., and Mouazen, A. M., "Wheat yield prediction using machine learning and advanced sensing techniques," *Computers and Electronics in Agriculture* **121**, 57–65 (2016).

[6] Ali, I., Cawkwell, F., Dwyer, E., and Green, S., "Modeling managed grassland biomass estimation by using multitemporal remote sensing dataa machine learning approach," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (2016).

[7] Akhand, K., Nizamuddin, M., Roytman, L., and Kogan, F., "Using remote sensing satellite data and artificial neural network for prediction of potato yield in bangladesh," in [*SPIE Optical Engineering+ Applications*], 997508–997508, International Society for Optics and Photonics (2016).

[8] Maresma, Á., Ariza, M., Martínez, E., Lloveras, J., and Martínez-Casasnovas, J. A., "Analysis of vegetation indices to determine nitrogen application and yield prediction in maize (zea mays l.) from a standard uav service," *Remote Sensing* **8**(12), 973 (2016).

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

[10] Ren, S., He, K., Girshick, R., and Sun, J., "Faster R-CNN: Towards real-time object detection with region proposal networks," in [*Advances in Neural Information Processing Systems (NIPS)*], (2015).

[11] Uijlings, J. R., Van De Sande, K. E., Gevers, T., and Smeulders, A. W., "Selective search for object recognition," *International journal of computer vision* **104**(2), 154–171 (2013).

[12] Girshick, R., Donahue, J., Darrell, T., and Malik, J., "Rich feature hierarchies for accurate object detection and semantic segmentation," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 580–587 (2014).

[13] Girshick, R., "Fast r-cnn," in [*Proceedings of the IEEE International Conference on Computer Vision*], 1440–1448 (2015).

[14] Ren, S., He, K., Girshick, R., and Sun, J., "Faster r-cnn: Towards real-time object detection with region proposal networks," in [*Advances in neural information processing systems*], 91–99 (2015).

[15] Zeiler, M. D. and Fergus, R., "Visualizing and understanding convolutional networks," in [*European conference on computer vision*], 818–833, Springer (2014).

[16] Tiebiao, Z., Jose, G., Jozef, F., Qi, Y., and YangQuan, C., "Melon classification and segmentation using low cost remote sensing data drones," in [*13th International Conference on Precision Agriculture*], The International Society of Precision Agriculture (2016).

[17] Yang, Q., Chen, D., Zhao, T., and Chen, Y., "Fractional calculus in image processing: a review," *Fractional Calculus and Applied Analysis* **19**(5), 1222–1249 (2016).