State Horticultural Association of Pennsylvania **Research Grant Proposal**

Title: Bloom Intensity Estimation using your Smartphone: Machine Learning Algorithms for Species-independent Visual Recognition of Flowers

Personnel:

Principal Investigator: Dr. Henry Medeiros, Assistant Professor, Marquette University, 1551 W Wisconsin Ave., Milwaukee, WI 53233. (414) 288-6186
henry.medeiros@marquette.edu

Co-principal investigator: Dr. Amy Tabb, Research Agricultural Engineer, USDA Agricultural Research Service (ARS) Appalachian Fruit Research Station (AFRS), 2217 Wiltshire Rd. Kearneysville, WV 25430. (304) 725-3451 x386
amy.tabb@ars.usda.gov

Duration of Project: One year. 9/1/2018-6/1/2019

Justification:

Need for bloom estimates in apple and peach production. Growers and extension specialists use bloom counts to schedule thinning sprays as a component of crop load management. In apple, the pollen tube model (Yoder et al. 2012) requires bloom counts as an input. The protocol for precision chemical thinning of apple in Robinson et al. (2014) also requires bloom counts. For peach, bloom estimates or densities have applications for precision crop load management, as well as automated thinning, such as with string thinners (Auxt Baugher 2010) or other selective means such as robotics (Lyons et al. 2015). The current practice in assessing bloom is human sampling.

Drawbacks of current practice. Human-performed sampling has many drawbacks. These drawbacks include, but are not limited to: 1) Tree inspection is time-consuming and the number of trees that can be inspected by the grower is limited due to this constraint; 2) Estimation of the blooming intensity or fruit load by visual inspection is characterized by a large number of uncertainties and is prone to errors; 3) Extrapolation of the results obtained by inspecting individual trees to estimates corresponding to entire rows or blocks relies heavily on the grower's skill and experience; 4) Inspection of a small number of trees does not provide information about the spatial variability which exists in the orchard; and 5) Experienced and skilled personnel to perform the sampling during a fast-moving spring may be limited or not available at the desired time. For these reasons, we pursue an automated method for estimating bloom counts from digital images.

Previous work on autonomous bloom estimation. Some previous works address the problem of autonomously assessing bloom counts. Some groups used customized data acquisition setups such as structured light and multispectral imaging. Nielsen et al. (2012) use a structured light system to detect and reconstruct flower positions in peach, while Wouters et al. (2015) use multispectral imaging and detect pear flower buds; both of these methods only operate at night. Horton et al. 2016 described a system using an unmanned aerial system (UAS) or drone for peach bloom intensity estimation. Based on the premise that the photosynthetic activity of this species increases during bloom period, the system relies on multispectral aerial images of the orchard, yielding an average detection rate of 84.3% for 20 test images. This method also has the intrinsic limitation of being sensitive to illumination conditions because of its use of hard thresholds.

Other research groups used color cameras as sensors for detecting bloom intensity. Underwood et al. (2016) detected flower pixels on almond trees and reported that simple thresholding operations in hue-saturation-value (HSV) space were sufficient to extract the flower pixels. However, whereas in almond trees flowers appear before leaves, in apple trees leaves and flowers appear at the same time, which renders the task of estimating blooming intensity much more challenging. In addition, in Underwood et al (2016) data was acquired in Australia, with a different color scheme than the eastern U.S. Hočevar et al. (2014)

followed a similar strategy, in that flower localization in apple trees is based on hard thresholding (in the hue-saturation-lightness color space) and size features, such that parameters must be adjusted whenever changes in illumination, flowering density (high/low concentration), or in camera position (near/far trees) occur. Blooming intensity of apple was investigated by Aggelopoulou et al. (2011), who acquired color images in broad daylight. Despite acquiring images at specified times and the presence of a black cloth screen behind the trees, due to distortion and oversaturation of images only 113 out of 250 images were usable. Aggelopoulou et al. (2011) did not have ground truth blooming data but rather calibrated a linear relationship between the number of white pixels and yield. Considering the large number of environmental and management factors which influence fruit development after the flowering stage, the robustness and usefulness of such a relationship would require further investigation. More recently, Krikeb et al. (2017) were able to quantify blooming intensity of individual apple trees using color images. Although they reported very high correlation between manual and automated estimation of blooming intensity (R^2>0.90), the relationships were season-specific and required local calibration, underscoring the need for further investigation.

Prior work by the team. The PI and co-PI recently devised a novel method for apple flower detection based on features extracted by a convolutional neural network (CNN) (Dias et al., 2018). In the proposed approach, flower region proposals are generated by grouping similar pixels into clusters known as superpixels (Achanta et al., 2012). Each region generated by the superpixel segmentation step is converted into a feature vector by a CNN, which is fine-tuned for saliency detection (Zhao et al., 2015). We further tuned this model for flower identification using labeled regions from a training set. Our feature vectors correspond to the output of the first fully connected layer of the network, which is provided as the input to a support vector machine (SVM) classifier, after dimensionality reduction. The SVM then determines whether each image region contains flowers or not.

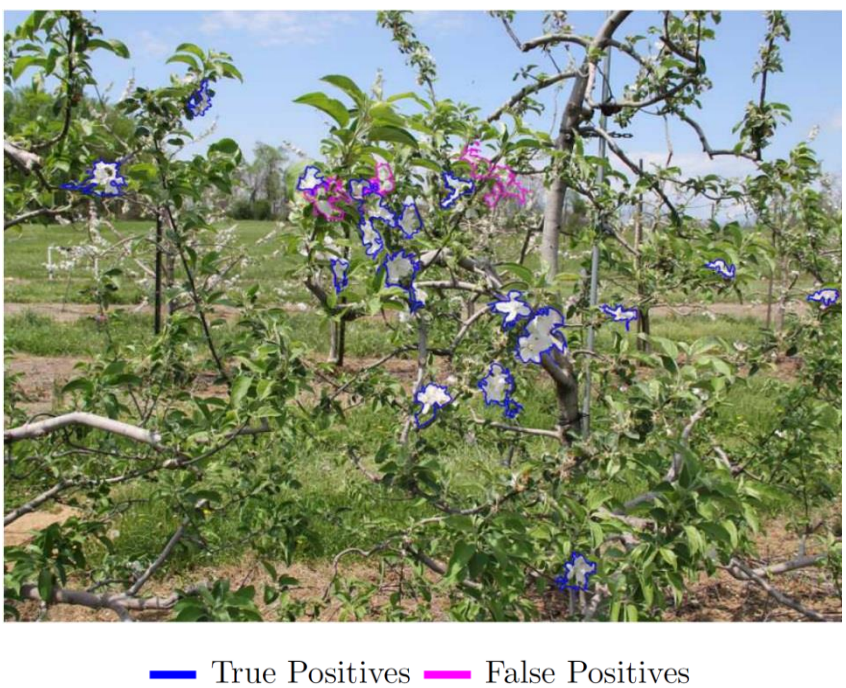

Figure 1. Example image, with result overlaid, from a dataset acquired by a consumer-level camera. The result of running our algorithm on this dataset is represented by colored outlines around regions: blue indicates a correct identification of a flower region and magenta indicates non-flower regions identified as flower regions by the algorithm.

Our algorithm was trained using a dataset of images of apple trees collected in a USDA-ARS-AFRS orchard in Kearneysville, WV under natural daylight illumination. The images in this dataset were collected by a hand-held Canon E0S 60D camera. A total of 147 images with resolution of 5184 by 3456 pixels were acquired under multiple angles and distances of capture. The dataset was manually labeled by indicating superpixels that contain parts of flowers in at least 50% of their areas. The images were randomly split into training and validation sets composed of a total of 91,488 training regions (i.e. superpixels) and 42,430 validation ones. See Figure 1 for an example of this dataset as well as some of the results from our method.

To evaluate the generalization capability of our method, we assessed its performance on our validation set as well as on three additional datasets. Two of the additional datasets also correspond to apple trees, but in these datasets a blue background panel is positioned behind the trees to visually separate them from trees in other rows of the orchard. The third additional dataset contains images acquired from a peach orchard during a cloudy day. Our experiments demonstrated that our approach shows optimal recall and precision rates higher than 90% for the validation set and near 80% for the all the other test datasets. A large number of superpixels classified as false positives correspond to regions where flowers are indeed present, but which were not labeled as positives since the flower corresponds to less than 50% of the total area of the superpixel. In other words, the sensitivity of the feature extractor to the presence of flowers is very high and the final performance would be improved if the region proposals were more accurate, which is part of the work proposed in this project.

Proposed system. We propose a low-cost sensing system for bloom estimation using consumer-grade color cameras such as those found in smartphones, hand-held digital cameras, or GoPros. Classification of bloom versus non-bloom regions is accomplished by a convolutional neural network (CNN) method, commonly referred to as deep learning. We propose implementing our algorithm in the form of a prototype web-based application (app). In this way, the technology can be used and evaluated by growers and extension personnel.

Our proposal fits in the **Ag Engineering** area, and covers all of the subject areas: *Sensor Technologies*, *Use of New Technology to Improve Data Collection for Decision Making*, and *Mechanization/ Automation Technologies for Improved Farm Efficiency*. In terms of *Sensor Technologies*, we will be using digital cameras in various forms for data acquisition. For the *Use of New Technology to Improve Data Collection for Decision Making*, one of the stated impacts of this work will be the ability to use an automated method to acquire bloom counts, which is needed as input for many aspects of decision making for crop load management: informing bloom thinning sprays, carbohydrate models, and pollen tube models. Finally, in the *Mechanization/ Automation Technologies for Improved Farm Efficiency* the bloom estimation will be done autonomously by computer vision systems and algorithms instead of by humans, leading to increased farm efficiency.

Objectives:

1. Strategy to increase regional detection accuracy of flower detection algorithms: investigate the applicability of end-to-end semantic segmentation convolutional neural networks to the problem of species-independent flower detection.
2. Strategy to create systems more robust to new orchard environments: Devise unsupervised or weakly supervised domain adaptation techniques to improve robustness to unseen background scenarios in fruit orchards.

Procedures:

Objective 1. In recent years, the combination of CNNs and increasingly larger image datasets has led to substantial improvements to the state of the art in image classification, object detection and semantic segmentation tasks. Semantic segmentation is the task of accurately localizing the contours of each object within an image and assigning category labels to these objects. Existing semantic segmentation datasets comprise tens of thousands of images containing up to 172 object categories (Caesar et al. 2016), and some of them include the category "flowers." Recall that in our current work, one of the sections identified as a cause of misidentifications of flower or non-flower regions was the regional grouping step referred to as superpixels. Since semantic segmentation methods tend to be robust to substantial variations in the appearance of objects within the same category, we intend to evaluate the performance of existing end-to-end convolutional networks as a starting point to segment fruit flowers and will compare the performance of the new approach with our existing approach that utilizes superpixels. Images composing publicly available datasets, however, typically contain only a few salient and large foreground objects, such that CNNs trained exclusively in these datasets have limited performance in real world applications. Hence, we intend to explore domain adaptation strategies that include incorporating a sliding window mechanism to deal with the fact that the flowers are relatively small compared to the image resolution. Moreover, we plan to fine-tune our algorithms using publicly available datasets of flower pictures as well as our own labeled datasets that focus particularly on fruit tree flowers.

Objective 2. Vision systems that rely on CNN approaches require large amounts of training data that is representative of the real-world conditions, and up to this point, all of this data in our group has been hand-labeled. Hand-labeling images that would represent the variety of conditions present for a robust flower detection system would be a daunting task; fruit flower appearance in images is not uniform. Different fruit flower species might present variations in color, size and shape. For different image datasets, the appearance of flowers from the same species might vary according to changes in illumination conditions, background composition and angle of capture. To improve the system robustness against such variations, we plan to increase the amount of training data using simulated images, a strategy being successfully used in the community to overcome the hand-labeling issue of CNNs (Rahnemoonfar and Sheppard, 2017; Sun and Saenko 2014). For instance, we plan to apply transformations such as rotation, scaling, warping and color space shifts to pre-labeled data to create new samples. Moreover, we plan to use a robotic arm in our laboratory to collect a large number of images of flowers and non-flower structures (e.g. branches, leaves) from different perspectives. We will position a highly contrastive background screen behind the flower/trees, such that synthetic images under a variety of environmental conditions can be easily generated by overlaying the images obtained with the background screen to real world background images. To further improve the system's adaptability, we will also investigate the effectiveness of a weakly supervised fine-tuning calibration procedure. Instead of requiring a large number of labeled images, this procedure would consist of fine-tuning the algorithms with a small set of site-specific images that comprise challenging scenarios, such as cluttered flowers, poor illumination and atypical structures composing the background.

Evaluation plan. Our methods will be assessed according to traditional machine learning metrics such as precision-recall curves, which summarize the trade-offs between the number of incorrect detections and the number of missed detections generated by an algorithm as its decision threshold varies. We will generate these metrics using our current annotated datasets of apple and peach flowers, which contain a total of 209 images in which 7,037 flowers were manually annotated, and compare the performance in terms of accuracy to our past work (Dias et al. 2018). Additionally, we will implement a data collection application for mobile phones that will be used to qualitatively validate the algorithms in the field and to collect additional data for training and testing our methods.

Budget:

Salaries $20,000	Supplies
Hourly wages	Travel
Fringe Benefits	Miscellaneous $200
Total $20,200	

Other Support:

Current support	USDA-ARS agreement with Marquette University #58-8080-5-020, amount of $58,195.
Pending support	Automated estimation of flower, fruitlet, and fruit load in apple orchards using computer vision. PIs. R. Linker, H. Medeiros, A. Tabb. Binational Agricultural Research and Development Fund (BARD). $302,000. ($150,000 to Medeiros and Tabb)

References:

1. R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk. (2012) SLIC superpixels compared to state-of-the-art superpixel methods, IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (11): 2274–2281.
2. Aggelopoulou, A., Bochtis, D., Fountas, S., Swain, K., Gemtos, T., Nanos, G. (2011). Yield prediction in apple orchards based on image processing. Precision Agriculture 12: 448-456.
3. Auxt Baugher, T., J. Schupp, K. Ellis, J. Remcheck, E. Winzeler, R. Duncan, S. Johnson, K. Lewis, G. Reighard, G. Henderson, M. Norton, A. Dhaddey, and P. Heinemann. (2010). String Blossom Thinner Designed for Variable Tree Forms Increases Crop Load Management Efficiency in Trials in Four United States Peach-growing Regions. HortTechnology 20(2): 409-414.
4. Caesar, H., Uijlings, J., & Ferrari, V. (2016). COCO-Stuff: Thing and Stuff Classes in Context. arXiv preprint arXiv:1612.03716.
5. Dias, P., A. Tabb, H. Medeiros. (2018). Apple Flower Detection using Deep Convolutional Networks. Computers and Industry [under review].
6. Hočevar, M., B. Širok, T. Godeša, M. Stopar. (2014). Flowering estimation in apple orchards by image analysis. Precision Agriculture 15 (4): 466–478. doi:10.1007/s11119-013-9341-6.
7. Horton, R., E. Cano, D. Bulanon, E. Fallahi. (2016). Peach Flower Monitoring Using Aerial Multispectral Imaging, 2016 ASABE International Meeting. doi:10.13031/aim.20162461520.
8. Krikeb, O., Alchanatis, V., Crane, O., Naor, A. (2017). Evaluation of apple flowering intensity using color image processing for tree specific chemical thinning. Advances in Animal Biosciences 8: 466-470.
9. Lyons, D.J., P.H. Heinemann, J.R. Schupp, T.A. Baugher, and J. Liu. (2015). Development of a selective automated blossom thinning system for peaches. Transactions of the ASABE 58(6):1447-1457.
10. Nielsen, M., D. C. Slaughter and C. Gliever. (2012). Vision-Based 3D Peach Tree Reconstruction for Automated Blossom Thinning, IEEE Transactions on Industrial Informatics, 8(1): 188-196, doi: 10.1109/TII.2011.2166780
11. Sun, B., & Saenko, K. (2014). From Virtual to Reality: Fast Adaptation of Virtual Object Detectors to Real Domains. In Proceedings of the British Machine Vision Conference.
12. Rahnemoonfar, M.; Sheppard, C. (2017). Deep Count: Fruit Counting Based on Deep Simulated Learning. Sensors, 17: 905.
13. Robinson, T. S. Hoying, M.M. Sazo, and Andrea Rufato. (2014). Precision Crop Load Management, Part 2. New York Fruit Quarterly. 22(1):9-13.
14. Wouters, Niels, Bart De Ketelaere, Tom Deckers, Josse De Baerdemaeker, Wouter Saeys, (2015) Multispectral detection of floral buds for automated thinning of pear, In Computers and Electronics in Agriculture, 113: 93-103, https://doi.org/10.1016/j.compag.2015.01.015.
15. Underwood, J. P., Hung, C., Whelan, B., & Sukkarieh, S. (2016). Mapping almond orchard canopy volume, flowers, fruit and yield using LiDAR and vision sensors. Computers and Electronics in Agriculture, 130: 83-96.
16. Yoder, K.S., G.M. Peck, L.D. Combs, and R.E. Byers. (2012). Using a pollen tube growth model to improve apple bloom thinning for organic production. II International Organic Fruit Symposium 1001. 207-2014.
17. Zhao, R., Ouyang, W.; Li, H.; Wang, X. (2015). Saliency Detection by Multi-Context Deep Learning. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 1265-1274.