ESTIMATION OF COTTON POPULATION DENSITY IN FIELD CONDITIONS USING DEEP CONVOLUTIONAL NEURAL NETWORKS

Yu Jiang
Changying Li
Andrew H. Paterson
University of Georgia
Athens, GA

Abstract

Population density is a major component of cotton fiber yield. It can be estimated by combining laboratory test of seed viability and sub-sampling of field germination, both of which are subjected to errors. The overall goal of this study was to develop a deep learning based approach to count germinated plants in field conditions using aerial images. An unmanned aerial system (UAS) with a consumer-grade color camera was used to image a cotton field containing 516 plots on the 19th day after planting. Collected color images were used to reconstruct an orthoimage of the entire field through commercial software. The orthoimage was partitioned to individual plots based on GPS information. Regions of interests (ROIs) were manually labeled in plot images, including 1952 plant regions and 309 weed regions. The labeled image regions were extracted and used to train and test a customized architecture of a convolutional neural network (CNN). Subsequently, plot images with region labels were used to train a Faster region-based CNN (Faster-RCNN) with the customized CNN architecture. The trained Faster-RCNN was used to identify and count cotton plants in each plot image. Simple linear regression analyses were conducted to evaluate plant counting accuracy. The customized CNN architecture achieved over 97% accuracy of classifying plant and weed image regions. Regression results showed that imaging-based plant counts were highly (R^2 = 0.92) correlated with manual counts. Both results suggested that the proposed approach could accurately count germinated plants in field conditions, providing useful information for yield estimation.

Introduction

Cotton is a crucial source of natural fiber and plays an important role in the United States economy (USDA-ERS, 2017). Increases in cotton fiber yield are not only significant to fulfilling fiber requirements of over nine billion people but also to farmers' profits. Population density is one of the most important yield components (Zhi et al., 2016). Traditionally, population density can be estimated using either laboratory germination tests or subsampled field assessments, but both are subjected to errors.

Recent advances in field-based high throughput phenotyping (FB-HTP) provide new tools for collecting various imaging data of plants, offering new opportunities to measure the number of germinated plants in field conditions (Li, Zhang, & Huang, 2014). In particular, the reduced cost and improved operability of consumer-grade unmanned aerial systems (UAVs) facilitate the research of calculating plant population density. Several studies have been conducted to use aerial color images to calculate the number of germinated plants (and therefore estimation of population density) (Chen, Chu, Landivar, Yang, & Maeda, 2017; Gnadinger & Schmidhalter, 2017). These studies relied on conventional image processing, which primarily counted plants based on segmentation results (counting by segmentation). This counting strategy has two major disadvantages. First, plants are usually segmented using thresholding methods, which is sensitive to plant greenness that can be significantly affected by illumination conditions and plant emergence statues. For instance, plants look darker on cloudy days than sunny days, and plants just sprout from the soil may have less greenness than well-established seedlings. Second, counting results are dependent on plant growth stages. Size filtering or regression models are usually used to address inaccurate counting of plants with overlaps: a large segmentation area can be counted as multiple plants based on plant size predetermined in a particular growth stage or estimated from regression models. Thus, it is necessary to adjust plant size filters or re-validate regression models for data collected from a new experiment field. In addition, breeding programs and genetics studies involve a wide variety of genotypes with a considerable variation in germination time, raising a particular concern of using these image processing based approaches for plant counting. In the past five years, convolutional neural networks (CNNs) have demonstrated a performance breakthrough in computer vision tasks such as object detection and recognition (He, Zhang, Ren, & Sun, 2016; Krizhevsky, Sutskever, & Hinton, 2017; Simonyan & Zisserman, 2014; Szegedy et al., 2015). In particular, Faster region-based CNN (Faster-RCNN) provided a general framework of using CNNs for various applications relied on object detection (Ren, He, Girshick, & Sun, 2015). Several studies utilized the Faster-RCNN framework for detection of crop fruit and weeds

(Kusumam, Krajník, Pearson, Duckett, & Cielniak, 2017; Potena, Nardi, & Pretto, 2016; Sa et al., 2017), and most of them showed a significant improvement in detection accuracy over methods based on conventional image processing. Therefore, it would be worthwhile to apply deep learning techniques (e.g. Faster-RCNN) for plant detection, evaluating a new counting strategy (counting by detection).

The overall goal of this study was to develop a deep learning based approach to count germinated plants in field conditions using aerial imaging. Specific objectives were to 1) collect and label image datasets for model training and testing, 2) train a Faster-RCNN with customized architecture for cotton plant identification, and 3) use the trained Faster-RCNN model to count germinated plants in individual plots.

Materials and Methods

Experimental field

A cotton field (33°43'10.9"N, 83°18'25.8"W) was planted in the Iron Horse Farm of the University of Georgia in Watkinsville, GA on 14 June 2017 (Figure 1). The field contained 516 plots (19 plots per row × 28 rows, excluding 16 empty plots) sampling 516 genotypes from two cotton populations. Plot and alley lengths and row-spacing were 3.05 m (10 feet), 1.83 m (6 feet), and 0.91 m (36 inches), respectively. A bag of 30 seeds were machine planted in each plot, with an average seed spacing of 0.1 m (4 inches). After planting, the GPS coordinates of start and end points were measured using a real time kinematic (RTK) GPS for individual plots, generating a GPS map for further data analyses. In addition, 14 ground control points (GCPs) were placed in the field and surveyed using the same RTK GPS.

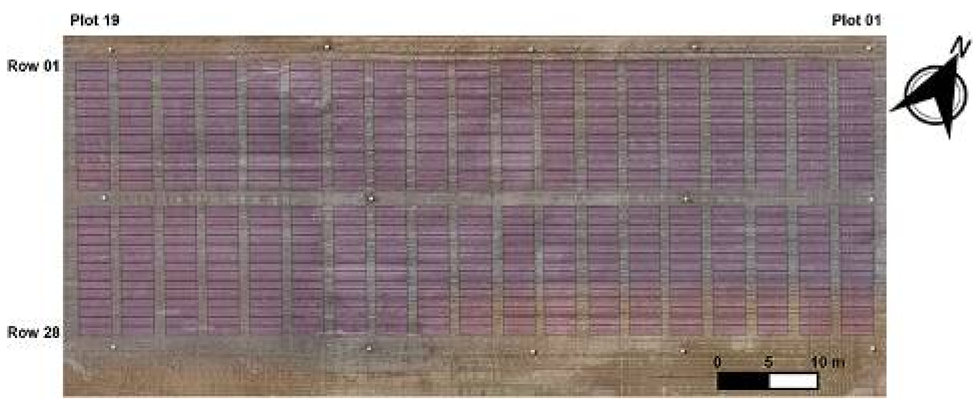


Figure 1. Orthoimage of the experimental field used in the study.

Data collection and preprocessing

A data pipeline was developed for the proposed deep learning based approach for cotton plant counting, covering from data collection to plant counting by detection (Figure 2). An unmanned aerial system (UAS; S1000+, Innovations Science and Technology Co., Ltd, Guangdong, China) with a consumer-grade color camera (Lumix DMC-G6, Panasonic Corporation, Osaka Prefecture, Japan) was used to image the field on 3 July 2017 when was the 19th day after planting (DAP 19). The day was mostly cloudy with an average wind speed of 1.34 m/s. The flight altitude was configured as 15 m above the ground. The camera was set to shutter priority mode, with a shutter speed of 1/800 s, a fixed focal length of 21 mm, and an ISO of 320. A total of 431 images were collected during the flight.

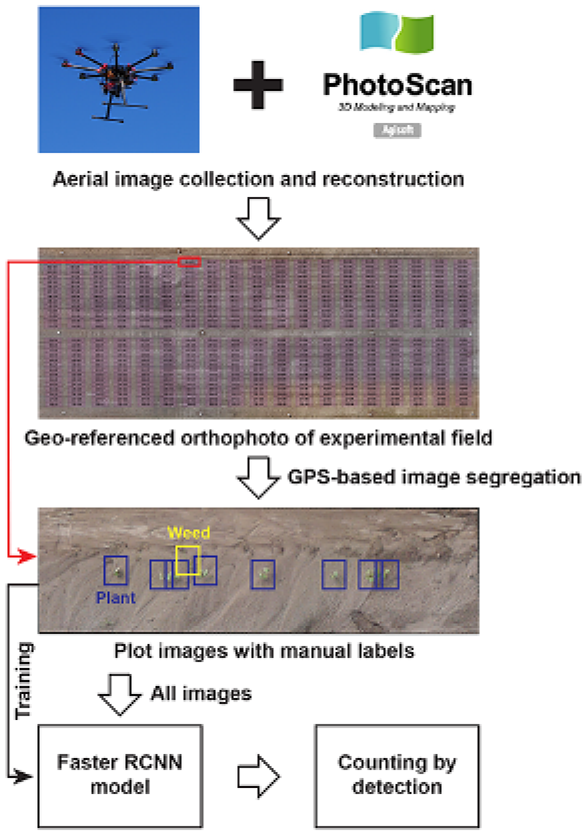


Figure 2. Data pipeline of the proposed deep learning based approach for cotton plant counting.

A commercial software (Photoscan Professional, Agisoft LLC, St. Petersburg, Russia) was used to generate an orthoimage of the entire field using collected aerial images. GCPs were identified in the orthoimage for geo-rectification. The orthoimage contained 39369 × 31842 pixels, with a ground sampling distance (GSD) of 2.72 mm/pixel. Based on the GPS map, the orthoimage was partitioned to individual plot images. Regions of interests (ROIs) were manually drawn for cotton plants and weeds in individual plot images, resulting in 2261 ROIs consisting of 1952 plant ROIs and 309 weed ROIs. It should be noted that ROIs were only drawn for regions containing a single cotton plant, because some plant regions might contain several plants with a significant overlap leading errors in network training. To train and test a CNN architecture, image patches of cotton plants and weeds were extracted from the labeled ROIs, and resized to the same size of 50 × 50 pixels.

Classifier training and testing

A Faster region-based convolutional neural network (Faster-RCNN) with a customized CNN architecture was trained in two steps. The first step was to evaluate of the customized CNN architecture using extracted image patches and the second one was to train a Faster-RCNN with the customized architecture using all plot images with labeled ROIs.

A customized CNN architecture was designed for differentiating seedling cotton plants from weeds (see Convolutional neural network in Figure 3). The CNN architecture contained four convolutional layers (conv layers) that were combined as two groups, with each group having a max-pooling layer afterwards. The filer size of all conv and max-pooling layers was 3 × 3 pixels, and the number of filters was initiated as 32 in the first conv layer and doubled after each conv layer. The conv layers used a stride step of 1 pixel with a padding of 1 pixel, reserving spatial information between consecutive conv layers within a convolutional group, whereas max-pooling layers used a stride step of 2 pixels with no padding, down-sampling images after a convolutional group. In the CNN architecture evaluation step, the feature map from the CNN was directly connected with two fully-connected layers associated with a classifier layer, providing class label for an input image patch. The extracted image patches were partitioned into three subsets (training, validation, and testing) with a ratio of 7:2:1. The training and validation sets were used for network training and parameter optimization, whereas the testing set was used only for evaluating the classification accuracy of the trained network. After evaluation of the CNN architecture, plot images with labeled ROIs were used to train the Faster-RCNN model which can localize and classify cotton plants simultaneously.

Training configuration was the same for the CNN and Faster-RCNN models. Momentum stochastic gradient decent (M-SGD) was used to optimize network parameters. The mini-batch size was 64 and the total epoch number was 20, resulting in 480 iterations in the training stage. The initial learning rate was set to 0.0001, with a decay of 0.1 for every 10 epochs.

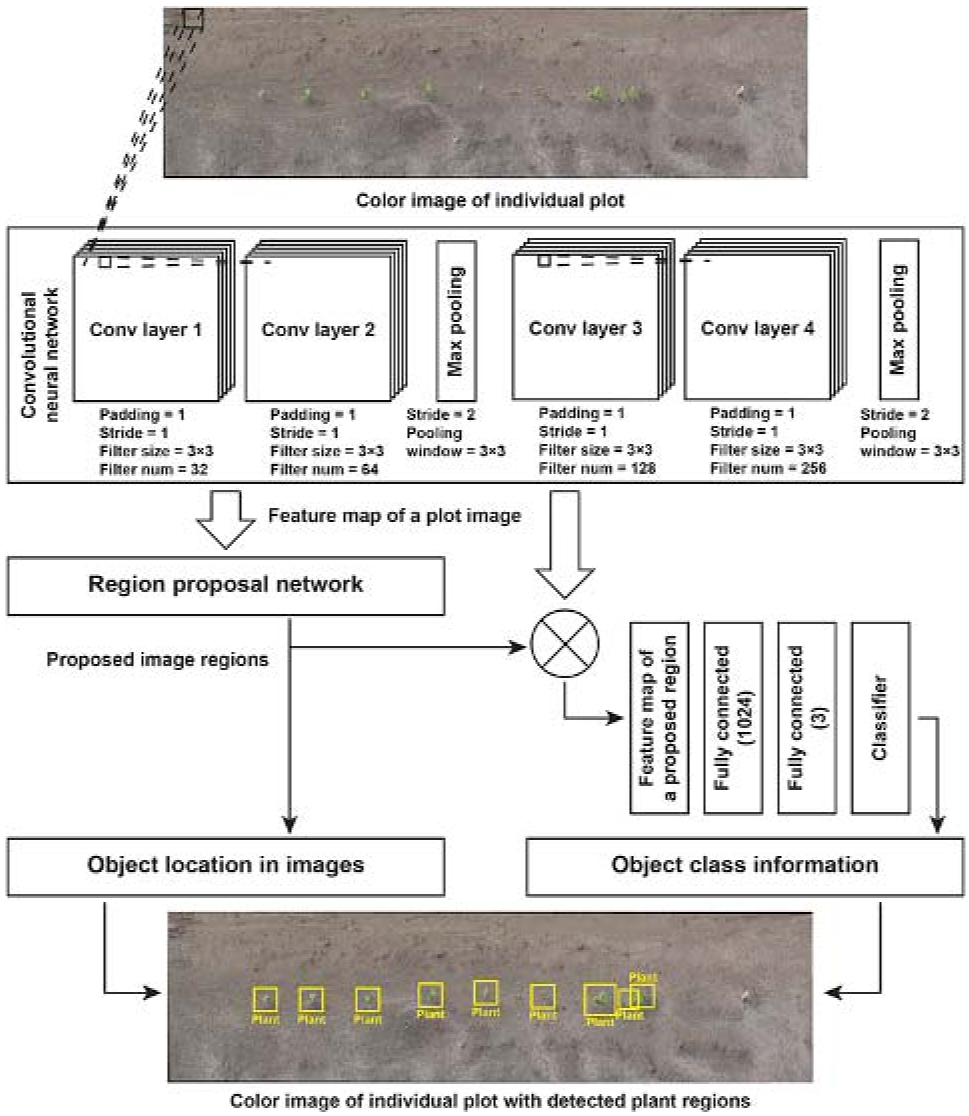


Figure 3. Architecture of the Faster region-based convolutional neural network (Faster-RCNN) used in the study.

Plant count by detection

The trained Faster-RCNN model was used to identify individual cotton plants in each plot image. The number of ROIs identified by the model was the number of plants counted by the proposed method. Since ROIs were labeled for a part of plants in each plot, it was less meaningful to directly compare model outputs and ground truth (labeled ROIs). Thus, it was necessary to compare the number of plants identified by the model and human field assessment. A total of 25 plots were randomly selected, and the number of plants in each plot was manually counted. Simple linear regression analyses were conducted between model and manual counts, and the coefficient of determination (R^2) and root mean squared error (RMSE) were used as indicator to evaluate the accuracy of the proposed method.

Results and Discussion

Example image patches of cotton plants and weeds

The experimental field contained a wide variety of genotypes having different emerging patterns (Figure 4). Some genotypes sprouted from the soil with two cotyledons (left column of plants), some started to form the first true leaf (middle column of plants), and some even established a couple of true leaves (right column of plants). The variation in emerging patterns caused an obvious difference in plant size. Although image patches were resized to the same size, the size of plants in the right column was the largest, followed by the middle and left columns. In addition, plants looked different in their appearance colors. If a plant sprouted for a short time period, its canopy could be covered with some soil, resulting in a reduction of contrast between the plant and ground. These two phenomena introduced challenges for traditional image processing based methods for plant counting, such as threshold and size based methods. In contrast, image patches of weeds showed less variation, because pre-emergency chemicals were applied to the field and only certain types of weeds could be germinated in the emergency stage.

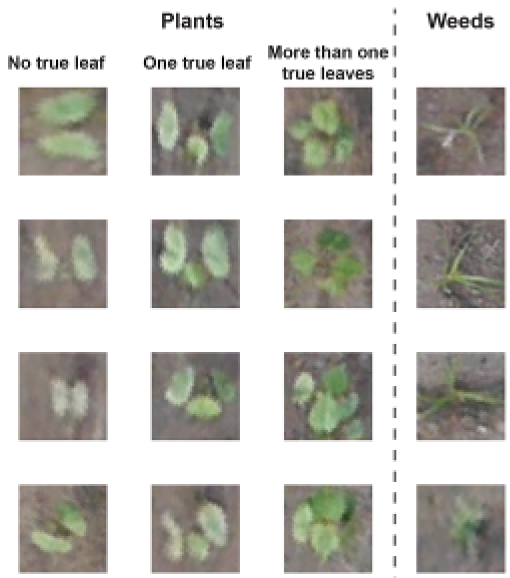


Figure 4. Example image patches of plants (with various emerging patterns) and weeds.

Classification accuracy of customized CNN architecture

The CNN model training can be grouped into fast converging (iteration 1 to 101), fine-tuning (iteration 102 to 230), and converged (after iteration 230) (Figure 5). In the fast converging stage, the mini-batch loss sharply decreased over iterations while the classification accuracy reached a local maximum accuracy (85%) after only several iterations. This occurred mainly because most plants and weeds were distinctive and filter parameters could be easily adjusted towards a local optimum that led to a fair classification accuracy. In the fine-tuning stage, the converging speed decreased over iterations until reached a local optimum in the 230th iteration. In this stage, filter parameters were further learned to differentiate plant samples that were similar to weed samples. Thus, despite of a slower converging speed, the classification accuracy started to increase dramatically. After the 230th iteration, the mini-batch loss and classification accuracy kept flat (with oscillation), indicating a performance plateau. Consequently, the 230th iteration could be considered as the converge point of the network training.

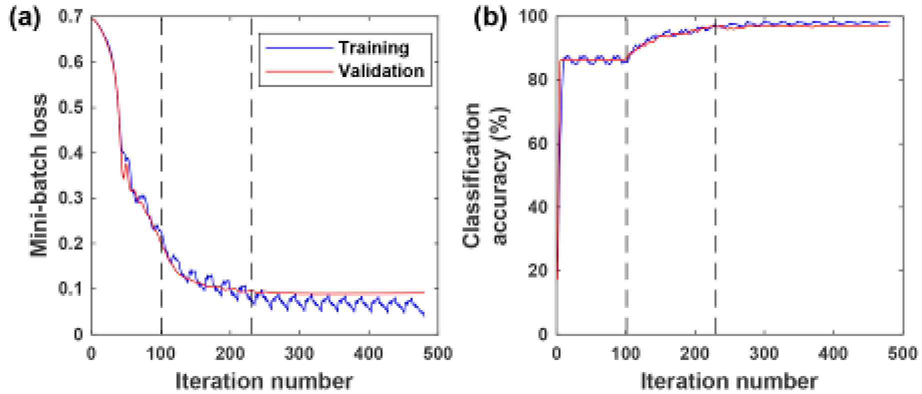


Figure 5. Mini-batch loss and classification accuracy of training and validation image sets.

There was no obvious gap in mini-batch loss and classification accuracy between the training and validation sets in the fast converging and fine-tuning stages, suggesting no overfitting in the two stages. On the contrary, the mini-batch loss and classification accuracy were further improved on the training set but not on the validation set in the converged stage, showing a potential of overfitting if the network kept being trained. Thus, the network trained in the 230th iteration was used as the final model for evaluating the performance of the proposed CNN architecture. The selected model achieved a classification accuracy of 97.32% on the testing set (Table 1), proving the efficacy of the proposed CNN architecture for classification of cotton plants and weeds.

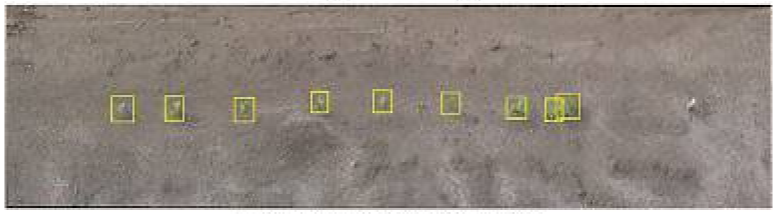
Table 1. Confusion matrix of classification of the testing images

	Labeled as plants	Labeled as weeds
Classified as plants	192	1
Classified as weeds	5	26
Accuracy	97.32%	

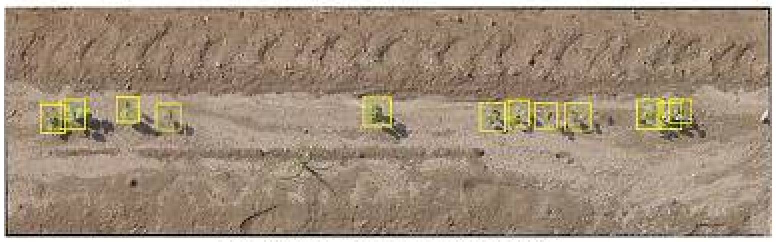
Accuracy of plant counting by detection

The trained Faster-RCNN model successfully identified individual cotton plants in plot images regardless of plot/plant and illumination conditions (Figure 6). There were three major challenges presented in the three images: 1) a low contrast between the cotton plant and the ground (see the top image in Figure 6); 2) various plant emerging patterns (see the middle image in Figure 6); and 3) overlap between cotton plants (see the bottom image in Figure 6). These three challenges were problematic for plant counting algorithms based on traditional image processing such as segmentation and morphological analyses of binary images, but they were overcome by the trained model. This was because the Faster-RCNN model (more specifically the CNN inside the Faster-RCNN) learned filters (feature extractors) based on ample information (e.g. color and edges) of an object in images. Hierarchical stack of the learned filters would allow the model to learn a combination of features that efficiently represented objects of a certain category, so the trained model could correctly identify plants with various emerging patterns (with/without true leaves) in different conditions (plant overlap and varying illumination).

Poorly germinated plot



Plot with well-separate plants



Plot with overlaps between plants

Figure 6. Representative results of plants identified by the trained Faster-RCNN model in plots with different germination conditions.

Regression analysis results showed that the number of plants counted by the proposed algorithm and human field assessment were highly (R^2 = 0.92) correlated with an RMSE of 1.61 (Figure 7), indicating a high accuracy of the proposed approach. However, it should be noted that the regression slope was 1.22, suggesting an underestimation of the number of plants counted by the approach. This underestimation was due to two potential error sources. First, aerial images were collected from nadir perspective, losing information from the side view, and thus it was impossible to image an entirely occluded plant. Second, the current camera resolution and dynamic range tended to be insufficient to resolve small plants, especially in areas where plant-ground contrast was low (see the top image in Figure 8). Nonetheless, some plots showed over estimation of plant counts (see the bottom image in Figure 8). The trained model incorrectly identified some weed regions as plant, resulting in an overestimation of plant counts. As most misclassified regions were between plots, they can be removed by using a location filter that preserves identified plant regions around the middle line of a plot, avoiding overestimation of plant counts. This needs to be further explored in future studies.

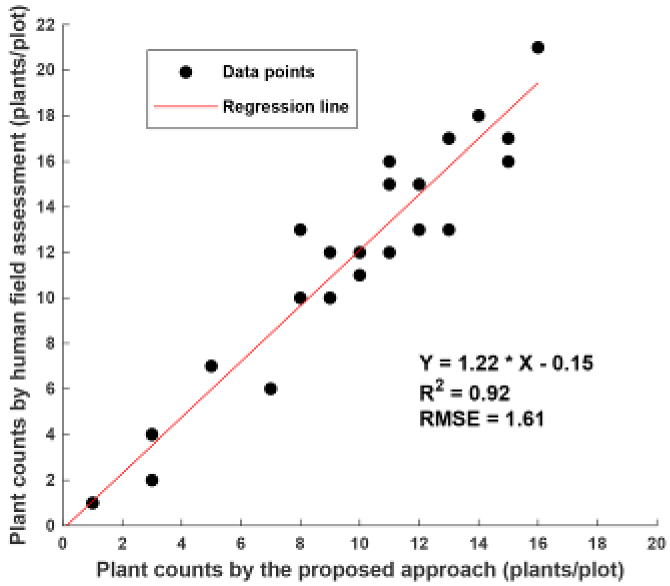


Figure 7. Regression results between plant counts calculated by the proposed approach and human field assessment for 25 randomly selected plots in the experimental field.

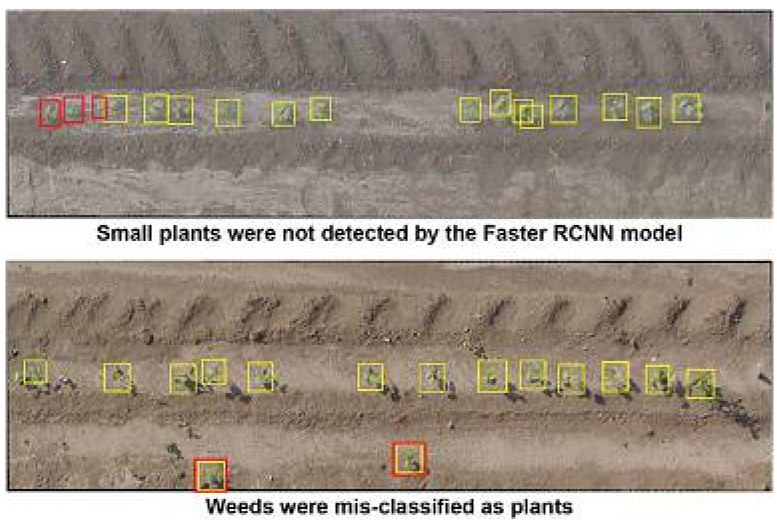


Figure 8. Potential error sources of plant counting: failure in plant detection (top image) and misclassification of plants (bottom image). Yellow rectangles indicate image regions identified as cotton plant in both images. Red rectangles indicate image regions of plant that were not detected in the top image, and misclassified image regions of plant in the bottom image.

Summary

The proposed CNN architecture achieved accuracies of over 97% on the training, validation, and testing sets after sufficient iterations, suggesting the efficacy of the proposed architecture for cotton plant classification. The trained Faster-RCNN based on the proposed CNN architecture could successfully identify individual cotton plants with various emerging patterns in different conditions. Regression results showed a high correlation (R^2 = 0.92) between the model-based and manual plant counts, indicating a high accuracy of the proposed method. Thus, the proposed approach could be an accurate tool for plant counting (and therefore population density calculation), providing useful information for cotton yield estimation for various application scenarios such as precision agriculture, breeding programs, and genetics studies. Future studies will be focused on improving the accuracy of identification of small plant in low contrast scenarios and reducing the overestimation caused by misclassification of plant regions.

Acknowledgements

This study was funded jointly by the Agricultural Sensing and Robotics Initiative of the College of Engineering, and the College of Agricultural and Environmental Sciences of the University of Georgia. The project was also partially supported by the National Robotics Initiative (NIFA grant No: 2017-67021-25928). The authors would gratefully thank Dr. Tariq Shehzad and Mrs. Jeevan Adhikari, Jon Robertson for field preparation and management, and Mr. Rui Xu and Shangpeng Sun for data collection.

References

Chen, R., Chu, T., Landivar, J. A., Yang, C., & Maeda, M. M. (2017). Monitoring cotton (Gossypium hirsutum L.) germination using ultrahigh-resolution UAS images. *Precision Agriculture*. doi:10.1007/s11119-017-9508-7

Gnadinger, F., & Schmidhalter, U. (2017). Digital Counts of Maize Plants by Unmanned Aerial Vehicles (UAVs). *Remote Sensing, 9*(6).

He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the Acm, 60*(6), 84-90.

Kusumam, K., Krajník, T., Pearson, S., Duckett, T., & Cielniak, G. (2017). 3D-vision based detection, localization, and sizing of broccoli heads in the field. *Journal of Field Robotics, 34*(8), 1505-1518.

Li, L., Zhang, Q., & Huang, D. F. (2014). A Review of Imaging Techniques for Plant Phenotyping. *Sensors, 14*(11), 20078-20111.

Potena, C., Nardi, D., & Pretto, A. (2016). *Fast and accurate crop and weed identification with summarized train sets for precision agriculture.* Paper presented at the International Conference on Intelligent Autonomous Systems.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). *Faster R-CNN: Towards real-time object detection with region proposal networks.* Paper presented at the Advances in neural information processing systems.

Sa, I., Lehnert, C., English, A., McCool, C., Dayoub, F., Upcroft, B., & Perez, T. (2017). Peduncle Detection of Sweet Pepper for Autonomous Crop Harvesting—Combined Color and 3-D Information. *IEEE Robotics and Automation Letters, 2*(2), 765-772.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., . . . Rabinovich, A. (2015). *Going deeper with convolutions.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

USDA-ERS. (2017). Cotton and wool: Overview. Retrieved from https://www.ers.usda.gov/topics/crops/cotton-wool/

Zhi, X. Y., Han, Y. C., Li, Y. B., Wang, G. P., Du, W. L., Li, X. X., . . . Feng, L. (2016). Effects of plant density on cotton yield components and quality. *Journal of Integrative Agriculture, 15*(7), 1469-1479.