INVESTIGATING RELATION BETWEEN OPEN BOLL COUNT AND COTTON SEEDLINGS EMERGENCE USING UNMANNED AERIAL SYSTEM IMAGES AND YOLO, A DEEP LEARNING FRAMEWORK

S. Oh
A. Chang
A. Ashapure
J. Jung
Texas A&M University—Corpus Christi
Corpus Christi, TX
M. Maeda
Texas A&M AgriLife Research
Lubbock, TX
J. Landivar
Texas A&M AgriLife Research
Corpus Christi, TX
S. Hague
W. Smith
Texas A&M University
College Station, TX

Abstract

Accurate estimation of cotton yield at early growth stages can make agricultural resource management more efficient throughout farming season. In this study, we try to estimate open boll count from number of cotton sprouts using unmanned aerial system (UAS) images and YOLO (You Only Look Once), a deep learning framework for real-time object detection. YOLO uses single convolutional network to predict the boundaries of existing objects and class probabilities simultaneously therefore it is suitable for large dataset as UAS images. The image data of this study were acquired at Texas A&M AgriLife Research and Extension Center at Corpus Christi, Texas, in April of 2017 when number of days after planting was less than 20. Total 466 images of 20 Megapixels were taken for 80 m by 170 m area where spatial resolution per pixel was about 0.55 cm. A YOLO network was trained from sub-dataset from original UAS images, and trained network model was used to locate center of each cotton seedling. Count of cotton seedlings and open bolls were aggregated using square grid along the rows of cotton plots. The correlation between number of cotton sprout and open boll is calculated and presented per each variety used in this trial.

Introduction

Plant phenotyping is assessment of plant traits such as growth, development, resistance, architecture, physiology and yield. With the application of unmanned aerial system (UAS) technology in agriculture, data collection and monitoring of cotton physiological properties are becoming more feasible. These physiological properties include canopy structure and vegetation indices which can be computed in relatively short period of time using simple equations. In addition to these properties, it is also becoming possible to detect shoot system such as leaves, buds, stems, flowers and fruits from UAS images by object detection and classification algorithms. Recently, many deep learning algorithms are available through variety of deep learning frameworks, and this allows us to identify plant organs with considerable detection rate. Detection rate and processing speed of deep learning algorithms is constantly improving, but most of deep learning algorithms were less applicable for UAS data due to the lack of performance in object detection and classification process compared to large amount of input data. However, speed of object detection was greatly reduced by YOLO (you only look once), a deep learning framework targeted for real-time applications. In this study, we aim to develop a technique that can detect cotton seedlings from UAS data using YOLO and find correlation between cotton seedlings and open boll count.

Methods

Study plot is located at Texas A&M AgriLife Research and Extension Center at Corpus Christi, TX. This area has humid, subtropical climate with hot, humid summer and mild, short winter. Cotton was planted on March 25, 2017 with 5 different cotton varieties, i.e., Gladdis, Tamcot73, T08, WK11L and X263, with solid and skip-row configuration where combinations of each variety with solid and skip-row were replicated 8 and 4 times, respectively. DJI Phantom 4 Pro® was used to collect UAS data that primarily comprises red, green blue (RGB) images embedded with geotagging information. The UAS platform used in this study is equipped with 1-inch 20-megapixel (5472 by 3648) sensor with maximum flight time of about 30 minutes. UAS data of the study area was collected on April 3, 2017, which is after 8 days after planting. A low altitude at 20 m with high side/front overlap of 85 % were used for the flight to obtain sampling distance of about 1 cm and redundancy for object detection. Subsequently, orthomosaic image was created using Agisoft Photoscan®, a structure from motion (SfM) based 3D reconstruction tool.

For object detection process, original images were first split to 608 by 608 square images resulting in about 25,000 in number. From entire set, 20 images were randomly chosen and divided into training and validation set with 3:1 ratio. Object boundaries of individual seedling were manually drawn by visual assessment from 15 training and 5 validation images. These training information were fed into Darknet-53, a convolutional neural network structure of YOLO, to train weights of the network, and accuracy of seedling detection was calculated by comparing results at each epoch and input validation data. Accuracy of seedling detection was maximized when iteration was 1,100 showing interest over union (IoU) was about 50 %. Using the network weights at this epoch, seedlings were detected from entire data set. Bounding boxes of all detected objects were projected onto reconstructed 3D ground surface, therefore multiple number of bounding boxes were present around individual seedlings. The multiple bounding boxes were then merged to single box by DBSCAN clustering algorithm.

Results and Discussion

Accuracy of deep learning based seedling count was measured by comparing it with visual counting from one of our researchers using UAS dataset. Twenty 2 meter strips were randomly selected within study plot and we compared the number of seedlings counted by two methods. As a consequence, the coefficient of determination of the two variables were 0.5255, and regression line was 1.235x – 6.503, which was similar to one-to-one line.

Assuming that the results of deep learning based seedling count is nearly accurate, we examined the correlation between seedling count and open boll count. From skip-row plots, average number of seedlings and open boll count per row was 92 and 1,208, respectively, and most of the value were scattered around the average without apparent correlation regardless of cotton variety. This indicates that density of cotton seedlings does not affect boll count when the varieties have enough spacing between rows. For solid-row plots, average number of seedlings and open boll count was 95 and 581, respectively. Comparing this with skip-row plots, number of cotton seedlings was nearly the same, however, open boll count was about the half lower in this case. Most data were centered around the average value, however, Gladdis, one of the 5 varieties in the experiment, exhibited coefficient of determination as 0.4196 between two variables. Although we cannot guarantee the correlation is statistically significant since only 8 replicates are available, this implies that we could roughly estimate open boll count of Gladdis at early growth stages by counting cotton seedlings.

Summary

A UAS-based cotton seeding counting procedure was suggested using YOLO deep learning framework. Correlation between seedling count measured by suggested method and visual assessment was 0.5255 and nearly followed one-to-one line. We applied this technique to investigate correlation between seedling and open boll count on a research plot located at Texas A&M AgriLife Research and Extension Center at Corpus Christi, TX. Among 5 varieties planted in the plot, Gladdis exhibited coefficient of determination over 0.4. In case more sample points are provided for each treatment condition, suggested methods can be used to roughly estimate open boll count from seedling count at early growth stage.