ABSTRACT

Author: Ashapure, Akash.
Title: Machine Learning based High-Throughput Phenotyping Framework for Crop Yield Prediction using Unmanned Aircraft Systems.
Major Professor: Dr. Jinha Jung

Estimating crop yield is of increasing importance to ensure agricultural stability, economic viability, and global food security. Provided with accurate crop yield estimation prior to harvest, farmers, breeders, and agriculture researchers are able to perform crop evaluation, genotype selection, and maximize yield by timely intervention. Remote sensing is often used to provide information about important canopy state variables for crop yield estimation. However, until recently, a key bottleneck in such research was the lack of high-throughput sensing technologies for effective and rapid evaluation of expressed phenotypes under field conditions for holistic data-driven decision making. Recent years have witnessed enormous growth in the application of unmanned aircraft systems (UAS) for precision agriculture. UAS has the potential to provide information on agricultural crops quantitatively, instantaneously, and, above all, nondestructively. This dissertation aims at utilizing UAS driven data to develop a machine learning based high-throughput phenotyping framework for crop yield estimation. In this research, plant parameters such as canopy height (CH), canopy cover (CC), canopy volume (CV), normalized difference vegetation index (NDVI), and excessive greenness index (ExG) were extracted from fine spatial resolution UAS based RGB and multispectral images collected on a weekly basis throughout the growing season. Initially, a comparative study was conducted to compare two management practices in cotton: conventional tillage (CT) and no-tillage (NT). This initial study was designed to test the reliability of the UAS derived plant parameters, and results revealed a significant difference in cotton growth under CT and NT. Unlike manual measurements, which relied on limited samples, UAS technology provided the capability to exploit the entire population. This makes UAS based data more robust and reliable.

Additionally, an inter-comparison study was designed to compare CC derived from RGB and multispectral data over multiple flights during the growing season of the cotton crop. A CC model that uses a multispectral sensor is considered more stable and accurate in the literature (Roth and Streit, 2018; Xu et al., 2019). In contrast, the RGB-based CC model is unstable and fails to identify canopy when cotton leaves change color after canopy maturation. With this study, it was

demonstrated that the application of morphological closing operation after the thresholding significantly improved the RGB-based CC modeling. The proposed RGB-based CC model provides an affordable alternative to the multispectral sensors that are more sensitive and expensive.

After assessing the reliability of UAS based canopy parameters, a novel machine learning framework was developed for cotton yield estimation using multi-temporal UAS data. The proposed machine learning model takes three types of crop features derived from UAS data to predict the yield. The three types of crop features are multi-temporal canopy features, non-temporal features (cotton boll count, boll size, boll volume), and irrigation status. The developed model provided a high coefficient of determination (R^2 ~ 0.9). Additionally, redundant features were removed using correlation analysis, and the relative significance of each input feature was determined using sensitivity analysis. Finally, an experiment was performed to investigate how early in the growing season the model can accurately predict yield. It was observed that even at 70 days after planting, the model predicted yield with reasonable accuracy (R^2 of 0.71 over test set). This study reveals that UAS derived multi-temporal data along with non-temporal and qualitative data can be combined within a machine learning framework to provide a reliable estimation of crop yield.

UAS technology is proven to be more robust and reliable. It efficiently works over small-size research fields or breeding trial fields. However, larger aerial coverage using UAS is not practically feasible. Alternatively, satellite images have the advantage of covering a vast area, but they provide coarser spatial resolution data. To overcome the limitation of UAS and satellite sensors, this study explored deep learning based methodologies to incorporate UAS based canopy attributes as additional information to improve the satellite-based yield estimation. It was accomplished through cross task knowledge transfer architecture and modality hallucination architecture. The main idea of this approach is to combine a multi-temporal satellite-based representation with an additional or complementary UAS based representation to improve crop yield estimation so that the models can predict the crop yield without utilizing UAS based representation at the test time. A significant improvement in the prediction accuracy was observed using cross task knowledge transfer and modality hallucination architecture. Additionally, the generalization capability of the proposed models was demonstrated by training on one experiment field and predicting crop yield for another field.