Nebraska Soybean Board
FINAL Research Report Form

VB 1/13/2020

Note: Submit this report no later than 90 days after the NSB-funded project officially terminates.
This post-project 90-day time-frame will allow the Lead PI time to complete any final data analysis and a final technical report, plus the drafting of any articles for submission to scientific journals. Note that this completed report will be provided to the National Soybean Checkoff Research Database, (soybeanresearchdata.com).

Project # and Title: #1721 Integrating high throughput field phenomics into Nebraska Soybean Breeding

Principal Investigator: Yufeng Ge

Co-PI's & Institutions: George Graef, University of Nebraska-Lincoln

Project Date (Including Extension): 10/01/2015 ***to*** 09/30/2019 ***(For example: mm/dd/yyyy to mm/dd/yyyy)***

Total Budget for Project: $ 165,000.00

1. Briefly State the Rational for the Research:

Plant breeding (and soybean breeding in particular for this research project) has entered an era where vast amount of genetic and genomic data become available to assist analysis and selection. The bottleneck is the availability of phenotyping data: Majority of the phenotyping efforts in a breeding program collects height (lodging), flowering time, staygreen in addition to final yield. The lack of high resolution phenotyping data limits the effective use of genomic data, as well as breeders' ability to discover new heritable traits for targeted crop improvement. This research project therefore had the rationale to build capacity in collecting and analyzing automated high-throughput field phenotyping data for Nebraska soybean breeding program. Together with other technologies such as genotyping and DNA sequencing, we hoped high-throughput phenotyping could speed up the breeding cycles and accelerate the rate of genetic gain in yield, quality and stress tolerance. In the long run, we hope these new technologies can be routinely implemented in soybean breeding practices and enhance the competitiveness of Nebraska growers in the international trade market.

2. Research Objectives: (copy from project, but keep in a brief bullet format)

The project had three specific research objectives.

Obj. 1. Create and develop an integrated sensing and engineering system with multiple crop and micro-climate sensors for low-cost and high-throughput field phenotyping data collection;

Obj. 2. Test the field phenotyping system and collect field phenotyping data in multiple years from multiple sites in Nebraska;

Obj. 3. Investigate image/data analysis pipelines to extract soybean traits from the raw sensor data and incorporate the field phenotyping data into the soybean breeding program.

3. General Approach Used and (if applicable) the Nebraska Test Locations:

For Obj. 1, we designed and developed a low-cost push cart that enabled semi-automated collection of plot-level field phenotyping data with high efficiency. The types of crop sensors included RGB cameras, ultrasonic sensors (for crop height measurement), infrared radiometers (for canopy temperature), and portable VIS-NIR spectrometers (for canopy reflectance from 400 to 1000 nm). The system also included two micro-climate sensors: a pyranometer (for solar radiation) and an air temperature/relative humidity sensor. A GPS receiver was included to provide georeferencing information for each plot. The cart was pushed manually through the field to position the sensor suites on top of each plot. The collection and storage of multi-faceted sensor data were automated via a custom-developed program. In normal field conditions, the cart could collect data from ~500 plots every hour. The ultrasonic sensors were later replaced by a 3D scanning LIDAR system, which could sample the height profile of the crop canopy at a finer resolution. Initially we used the portable spectrometers from Thorlabs. They were later replaced with the spectrometers from OceanOptics, which was more robust for field applications.

The push cart system was tested at various locations in Nebraska in 2016-2019 (four growing seasons). These locations included: Lincoln, Mead (Eastern Nebraska Research and Extension Center), Clay Center, Wymore, Cotesfield, and Valley. These field plots belonged to different studies including drought (water use efficiency) study, nitrogen study, germplasm evaluation, and iron deficiency chlorosis. For most of the plots, data collection was conducted twice in a season: first at V4/V5 stage and second at R5 stage. For some plots in Lincoln, data collection was conducted at multiple (6~7) times in a season. The total number of plots scanned with the push cart was > 20,000 over the four year period.

Algorithms and data analysis pipelines were developed to process the raw phenotyping data and extract useful traits. The focus was on (1) the processing of plot-level RGB images to extract vegetation coverage and leaf color variation; (2) the processing of canopy reflectance to extract useful spectral indices such as NDVI, Photochemical Reflectance Index, and chlorophyll index; and (3) the processing of 3D LIDAR point clouds to extract plot-level canopy height. Canopy temperature by the infrared radiometer was used together with the air temperature measurements to calculate canopy temperature depression. All extracted trait data were linked to the plots via their GPS coordinates.

4. Describe Deliverables & Significance Attained for Each Research Objective:

We developed a low-cost, semi-automated phenotyping platform that enabled the collection of multi-faceted, high throughput field phenotyping data from soybean breeding plots. This sensor platform was extensively tested in four growing seasons and over wide geographic regions in Nebraska. This sensor platform was proven to be very robust and reliable in field applications. This platform was also used by Graduate Student Shawn Jenkins to collect soybean phentoyping data from other states in 2019 (MO, KS, IA, etc.). In addition to soybean, this platform was also used to collect field phenotyping data for other crops including wheat, camelina, and early-season maize and sorghum (therefore impacted a wider research community at University of Nebraska-Lincoln).

One unique research opportunity for the team was the work with soybean Iron Deficiency Chlorosis (IDC). IDC rating was currently done by visual scoring and suffered from low inter-rater repeatability. We hypothesized that plot-scale RGB images captured by the field phenotyping platform could be used to extract leaf color information and develop a more automated and objective approach for soybean IDC scoring. Therefore in 2016 and 2017, the team conducted field experiments to collect images from IDC trials at multiple locations. Image analysis algorithms were developed to extract color/size features from soybean plot images (at V3/V4 stages). These features were then used to train machine learning models (Linear Discriminant Analysis and Support Vector Machine) to classify different soybean genotypes into IDC score categories. We showed that this approach based on field phenotyping images performed slightly more accurate than human raters for IDC scoring. Inspired by this research, further studies to employ more sophisticated convolutional neural network and deep learning for IDC scoring were undergoing. Automated soybean IDC scoring using Unmanned Aircraft Systems was also conducted and investigated. Taken together, these projects aimed at developing automated, more accurate/repeatable IDC scoring methods for the development of more IDC resistant soybean cultivars.

Nebraska Soybean Board
FINAL Research Report Form

4. Describe Deliverables & Significance Attained for Each Research Objective (continued)

Another analysis we did with the multiple-year, multiple-location phenotyping data was the prediction of soybean yield traits (including yield, maturity, height, seed size, protein, oil, fiber, lodging, and seed quality) from the plot-level RGB images taken at early season (V4/V5 stages). We used over 6000 plot-level RGB images and extracted over 450 color and textural image features. These image features were used to develop statistical models to predict the final yield traits. The results showed that many of these traits could be predicted from early season imagery with moderate accuracy (R2 > 0.50), suggesting a potentially useful strategy of predicting end-of-season yield traits from early season imagery.

In this project, we collected over 1TB of soybean phenotyping data. While we already did substantial analyses with the datasets, it should be noted that there are many other ways that these phenotyping datasets can be analyzed for. In particular, when the genotyping information for these soybean lines become available, they would enable other analyses such as QTL mapping, GWAS, or Genomic Prediction. Some of these analyses have already been conducted by Dr. George Graef's research group. These phenotyping datasets would also be available to the research communities outside the UNL.

This project trained one postdoctoral research associate (Dr. Frank Bai). In April 2018, Dr. Bai was promoted as Research Assistant Professor in charge of UNL's Spidercam field phenotyping facility. The project also trained one MS student (Dr. Wenan Yuan). In Aug 2019, Wenan successfully defended his thesis and started to pursue his PhD at Penn State University. Dr. George Graef's PhD student Shawn Jenkins had worked closely with my group in this project. Shawn will defend his dissertation in Feb 2020. Data generated from this project contribute significantly to Shawn's research.

Finally, this project allowed the team gather a lot of data and experience to attract extramural funding in high-throughput plant phenotyping research. Some examples of the funding are listed below.
NSF IDBR: TYPE A: Multispectral laser 3D ranging and imaging system for plant phenotyping. $534k.
USDA-NIFA PAPM EAGER: Transitioning to the next generation plant phenotyping robot. $285k.
USDA-NIFA Genomics and phenomics to identify yield and drought tolerance alleles for improvement of camelina as biofuel crop. $1M.
North Central Sun Grant Accelerating improvement of biomass sorghum as biofuel feedstock via high throughput phenotyping. $200k.

5. List where the Project Research Results/Findings were Publicized:

This project resulted in six peer-review journal publications. NSB funding was acknowledged in all these publications.

Bai, G., Ge, Y., Hussain, W., Baenziger, P.S., Graef, G., 2016. A multi-sensor system for high throughput field phenotyping in soybean and wheat breeding. Computers and Electronics in Agriculture 128, 181-192. https://doi.org/10.1016/j.compag.2016.08.021

Pandey, P., Ge, Y., Stoerger, V., Schnable, J.C., 2017. High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. Frontiers in Plant Science 8, 1348. https://doi.org/10.3389/fpls.2017.01348

Bai, G., Jenkins, S., Yuan, W., Graef, G., Ge, Y., 2018. Field-based scoring of soybean iron deficiency chlorosis using RGB imaging and statistical learning. Frontiers in Plant Science 9, 1002. https://doi.org/10.3389/fpls.2018.01002

Yuan, W., Li, J., Bhatta, M., Shi, Y., Baenziger, P., Ge, Y., 2018. Wheat height estimation using LiDAR in comparison to ultrasonic sensor and UAS. Sensors 18(11), 3731. https://doi.org/10.3390/s18113731

Yuan, W., Wijewardane, N.K., Jenkins, S., Bai, G., Ge, Y., Graef, G.L., 2019. Early prediction of soybean traits through color and texture features of canopy RGB imagery. Scientific Reports 9, 14089. https://doi.org/10.1038/s41598-019-50480-x

Bai, G., Ge, Y., Scoby, D., Leavitt, B., Stoerger, V., Kirchgessner, N., Irmak, S., Graef, G., Schnable, J., Awada, T.N., 2019. NU-Spidercam: A large-scale, cable-driven, integrated sensing and robotic system for advanced phenotyping, remote sensing, and agronomic research. Computers and Electronics in Agriculture 160, 71-81. https://doi.org/10.1016/j.compag.2019.03.009

Note: The above boxes will automatically accomodate for your text inputs; HOWEVER, the Final Report comprised of the above listed items must be kept to THREE PAGES. A Technical Report of no more than TEN PAGES (preferably fewer) can be appended to this report.
Submit both reports as a single PDF with this file name format: #XXX > FINAL > Project Title > PI last name

Please email this completed form to the Agriculture Research Division (jmonaghan2@unl.edu) based on the reporting schedule given to you. If you have any questions, please call the ARD at 2-2045 or Victor Bohuslavsky at the Nebraska Soybean Board Office at (402) 432-5720.