Project report (Second quarter April 1 2023 to June 30, 2023)

Project funded by North Central Soybean Research Program

Project tile - Field phenotyping using machine learning tools integrated with genetic mapping to address heat and drought induced flower abortion in soybean

Participating institutions – Texas Tech University, Kansas State University, University of Missouri, and University of Tennessee

Goals & Objectives

Long-term Goal – Develop soybean cultivars with 20 to 30% lower flower abortion under favorable to challenging environmental conditions, leading to about 10-15% increase in yield potential

Objectives (Year 1)

- Explore the genetic diversity in flower abortion under different soil moisture and climatic conditions using a large diversity panel

- Develop an image-based field phenotyping system and deep-learning tool to precisely document temporal dynamics in flower abortion and pod retention in genetically diverse soybeans

- Discover environmentally stable and region-specific genomic regions controlling flower abortion in diverse soil types, moisture, and climatic conditions

Progress achieved

Objective 1 - Explore the genetic diversity in flower abortion under different soil moisture and climatic conditions using a large diversity panel

Texas Tech University

Seed processing and field preparation activities were initiated for the 228 lines on May 19th, following the delivery of seeds on May 18th, from University of Missouri. On June 13th, the seeds underwent treatment with Histick (Basf) - Inoculant and Biofungicide to promote seed emergence, growth and protection.

However, unfavorable weather conditions characterized by frequent showers posed challenges, resulting in a delay in planting. On June 16th, the soybean seeds were planted, and the plants have currently progressed to the V3 growth stage (Figure 1).

Figure 1. Soybean plants in stages V2 - V3.

To ensure effective weed control, continuous monitoring efforts have been undertaken in the field. To facilitate image phenotyping, our team is currently exploring the use of a sprayer or a tractor with a sprayer implement for installing the cameras (Figure 2).

Figure 2. Sprayer and tractor sprayer implement options for installing the cameras.

University of Missouri

Seeds of a diverse set of soybean germplasm (228 lines) in the USDA Gene Bank were successfully increased in Costa Rica. Our group distributed seeds to collaborators in Tennessee, Kansas, and Texas in May.

We planted all these entries in Columbia, MO on May 24, 2023. Germination was excellent. Field plots are well established, and plants reached V4-V5 growth stages as of June 30th (Figure 3). We expect initial flowering in 10-14 days.

Figure 3. Soybean plants in stages V4 - V5.

We are preparing image-based field phenotyping system as instructed by the engineering group in this project and field phenotyping is expected to start in last week of July or first week of August.

University of Tennessee

Plots were planted on June 7, 2023 at WTREC. All 700 plots are well maintained. The beans are at growth stage V3 to V4 (3 to 4 trifoliate leaves) (Figure 4). The soybean crop will be managed according to University of Tennessee recommendations for growth regulator, pesticide applications, etc.

Figure 4. Soybean plants in stages V3 - V4.

Rainfall and environmental data will be provided by the National Oceanic and Atmospheric Administration Global Historical Climatology Network Weather Station (GHCND: USC00404561) located at the immediately adjacent the experimental field. A Ph.D. student is on board with us to start his dissertation research actives on the current soybean project.

Kansas State University

The planting of the soybean plots took place on May 25th. Currently, we are actively monitoring the plots, and it is anticipated that the soybean plants will soon reach the R1 growth stage (Figure 5).

Figure 5. Soybean plants reaching stage R1 soon.

To facilitate the installation of the imaging system, we have made specific modifications to a high-clearance spray vehicle (Figure 6). The wheel spacing has been adjusted to straddle our 10' wide plots, which will serve as the mounting platform for the imaging system. This modification ensures optimal coverage and accessibility for capturing high-quality images of the soybean plants.

Figure 6. High-clearance spray vehicle.

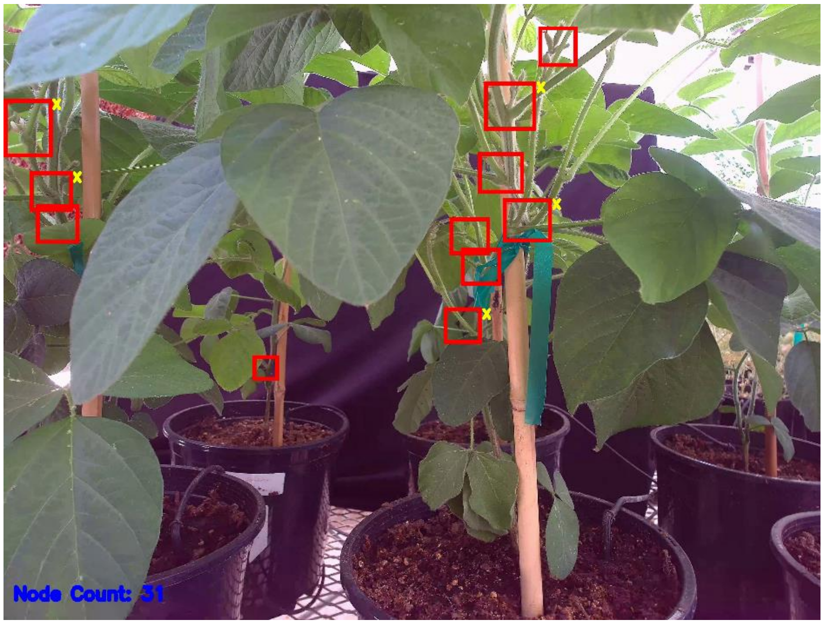
Objective 2 - Develop an image-based field phenotyping system and deep-learning tool to precisely document temporal dynamics in flower abortion and pod retention in genetically diverse soybeans

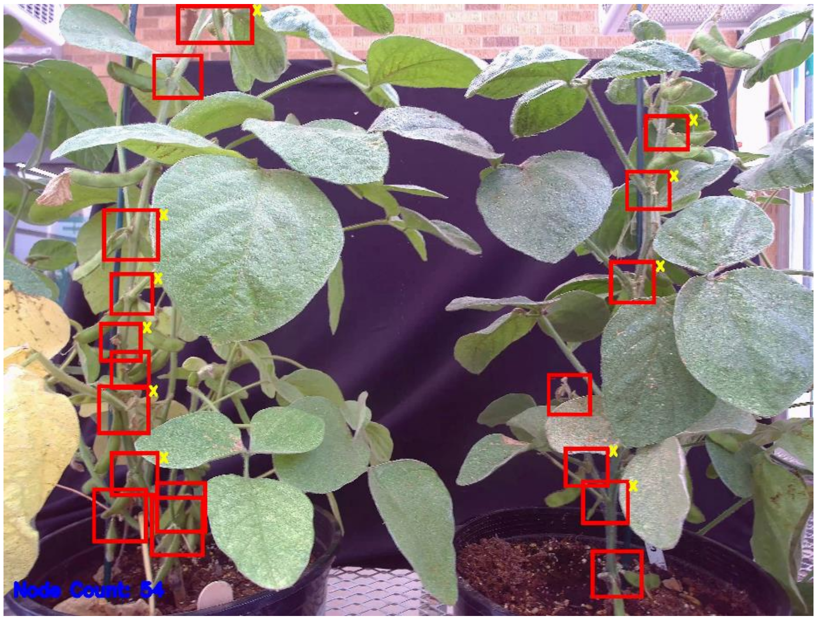
Texas Tech University

Five different models for node detection were evaluated, all of which were found to have comparable performances. We will hand these over to the K-State team so that they can integrate them into the flower detection pipeline and begin to process the images that will be collected at various sites in the coming weeks.

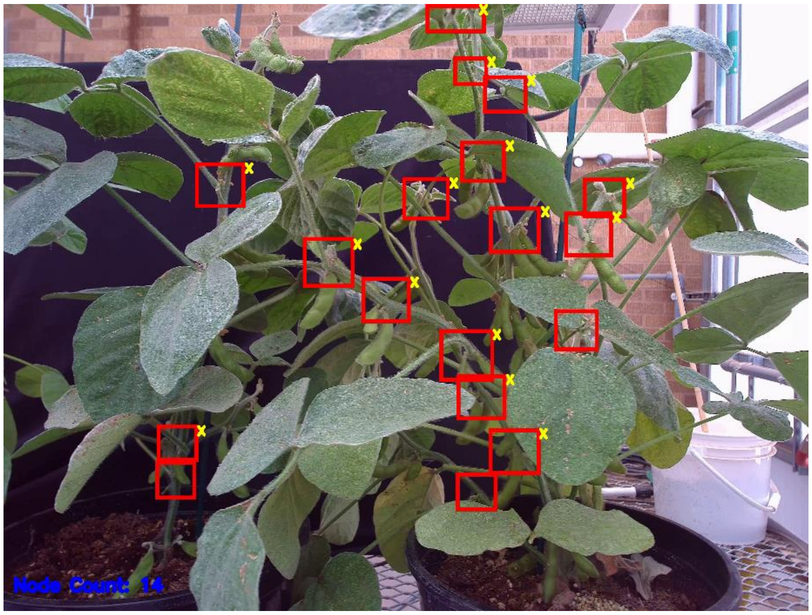
A GoPro Hero11 camera 27-megapixel was evaluated due to easy use and image collection. A protocol for image quality collection was developed based on the GoPro camera parameters to all locations. The imaging system was tested in a greenhouse and its ability to capture and record high-quality images at 60 frames per second was verified (Figure 7). Furthermore, the captured images were used as input to the node detection model with successful outcome (Figure 8).

Figure 7. Imaging system test in a greenhouse to set camera parameters and check the ability to capture and record high-quality images at 60 frames per second.





Node Count: 54



Node Count: 14

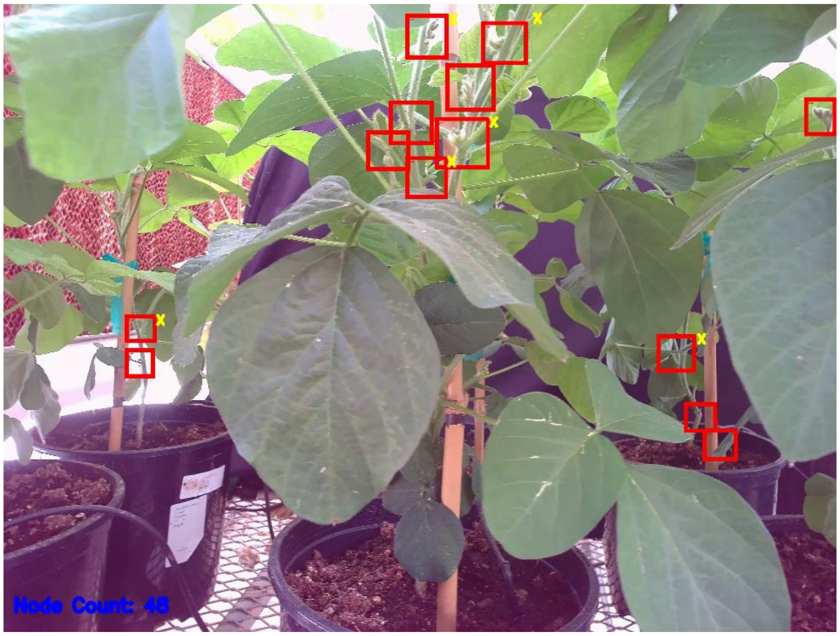


Figure 8. Soybean node detection.

We expect that the respective teams in each of the location will innovate, assemble and implement a strategy for conveying the imaging system through the field. We will provide back-stopping and help with image processing as the teams start generating field-imaging videos.

Kansas State University

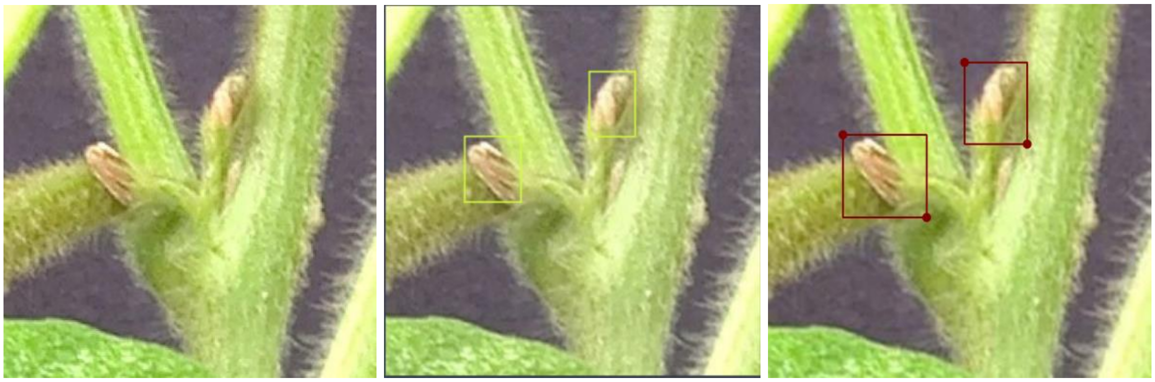
Improving the quality of the flower detection model

We have fine-tuned the original Faster R-CNN flower detection model to improve its predictions. Specifically, the model was fine-tuned with a variety of images, some taken in a more controlled environments and others resembling images taken in-the-field; some more focused, and others somewhat blurred; or images taken with different imaging systems/cameras producing different resolutions and quality. The Average Precision for detections whose bounding boxes overlap by at least 50% with the ground truth bounding boxes (denoted as AP50) was 79.53 on the test images. Some sample predictions on test images are provided (see PDF attachment), together with their corresponding ground truth annotations (the predicted and ground truth counts are also shown underneath each image) (Figure 9).

Node 3 IMX519

Original Image *Predicted 5* *Ground truth 4*

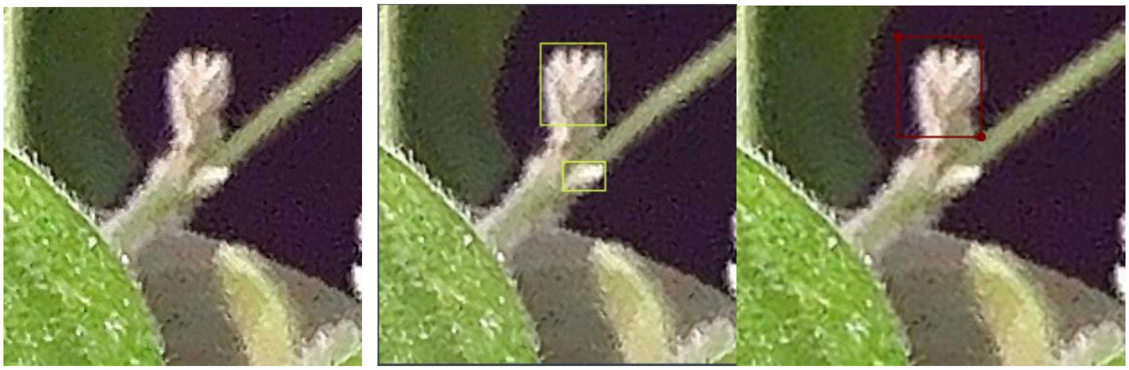
Node1 IMX519



Original Image *Predicted 2* *Ground truth 2*

IMX 477

Node 15



Original Image *Predicted 2* *Ground truth 1*

Node 13

Original Image *Predicted 0* *Ground Truth 0*

Figure 9. Prediction and ground truth counts.

Adding pods to the flower model

During the last reporting period, we have also enriched our model with the ability to detect pods. Specifically, we have adapted the previous Faster R-CNN model to detect pods (in addition to flowers) by fine-tuning it with 2693 annotated pod images (Table 1). We used the Faster R-CNN implementation available in Detectron2 (a library containing state-of-the-art detection and segmentation algorithms made publicly available by Facebook AI Research).

Table 1. Annotated pod numbers.

Subset	Number of images
Training	1889
Validation	537
Testing	267
Total	2693

In Figure 10 are predicted bounding boxes by comparison with the ground truth annotations along with the original images

Ground Truth 4 pods

Predicted 4 Pods

Original Image

Ground Truth 1 pod

Predicted 1 Pod

Original Image

Ground Truth 5 pods

Predicted 4 Pod

Original Image

Ground Truth 3 pods Predicted 2 Pod Original Image

Figure 10. Predicted bounding boxes by comparison with the ground truth annotations along with the original images.

Flowers/Pods per whole plant images

To better estimate the overall prediction capability of the flower/pod detection model, we evaluated it by comparing the number of detected flowers/pods with the number of ground truth flowers/pods per whole plant image (Note that this is different from the number of flowers/pods per plant, as some flowers/pods may not be visible in a particular image, depending on the angle of the image.) More specifically, we mapped the coordinates of the flowers/pods in each individual node image to coordinates in the whole plant image. This allows us to avoid duplicate detections. Some examples of predictions for each node in a plant image are shown in Figures 11 and 12.

Figure 11. Mapping the coordinates of the flowers/pods in each individual node image to coordinates in the whole plant image.


Node1 6 Flowers


Node 2 2 Flowers


Node 3 3 Flowers


Node 4 2 Flowers 1 Pod


Node 5 2 Flowers 2 Pods


Node 6 3 Flowers 1 Pod


Node 7 2 Flowers


Node 8 2 Flowers 1 Pods


Node 9 3 Flowers


Node 10 3 Flowers 1 Pod

Figure 12. Flower and Node detection.

The number of flowers detected from this plant from this angle is 28 flowers and the number of detected pods is 6.

Objective 3 - Discover environmentally stable and region-specific genomic regions controlling flower abortion in diverse soil types, moisture, and climatic conditions

Texas Tech University

Floral organ abscission is an important process that regulates the detachment of flowers from the stem. Floral organ abscission in well characterized model species (Arabidopsis) involves four steps: Initiation of abscission zone (AZ), promotion of AZ by ethylene, activation of separation and deposition of protective layer where organs have detached from the plant. We have shortlisted 6 genes (Blade on Petiole (BOP), KNAT (KNOX genes), BREVIPEDICELLUS 1 (BP1), INFLORESCENCE DEFICIENT IN ABSCISSION (IDA), HAE/HSL (leucine-rich repeat receptor like kinase), and DNA BINDING WITH ONE FINGER 4.7 (DOF4.7)) in Arabidopsis which correspond to 27 orthologous genes in Soybean involved in floral organ abscission. In addition, we shortlisted additional genes which have been reported to also play a role in floral organ abscission in addition to their known function- ASYMMETRIC LEAVES1 (AS1), AGAMOUS-like 15 (AGL15), and FOREVER YOUNG FLOWER (FYF). The mutant alleles of these genes have shown significant effect on several stages of floral organ abscission. And lastly, the maturity locus E1-E4 plays a significant role in the regulation of flowering in soybeans. The J locus, ortholog of AtELF3 (EARLY FLOWERING 3), is under the influence of E1. The functional analysis of mutant alleles for these genes showed an early flowering phenotype. The haplotype analysis for these genes is currently in progress, the analysis shows that some higher maturity group (MG) lines used in the current project (MG III, IV) retain one or more of the variant alleles. From this analysis, a group of lines correlating large effect variants associated with flowering traits (floral initiation and flower abortion) will be selected to identify causal genomic regions and thereby underlying genes.