

Ashlyn Rairdin¹, Fateme Fotouhi¹, Jiaoping Zhang¹, Daren S Mueller¹, Baskar Ganapathysubramanian¹, Asheesh K Singh¹, Somak Dutta¹, Soumik Sarkar¹, and Arti Singh¹

¹Iowa State University

October 5, 2022



NAPPN Annual Conference Abstract: Genome wide association study of sudden death syndrome in soybean using deep learning-based phenotyping

Ashlyn Rairdin¹, Fateme Fotouhi¹, Jiaoping Zhang¹, Daren S. Mueller¹, Baskar Ganapathysubramanian¹, Asheesh K. Singh¹, Somak Dutta¹, Soumik Sarkar¹, Arti Singh¹
¹Iowa State University, Ames, Iowa, United States

ORCID: 0000-0003-1544-6227

Keywords: stress phenotyping, disease quantification, object detection, machine learning, foreground detection, ROC analysis, image-based phenotyping, sudden death syndrome

Reliable and accurate method to phenotype disease incidence and severity is essential to unravel the complex genetic architecture of disease resistance in plants, and to develop disease resistant varieties. Genome-wide association studies (GWAS) involve phenotyping large numbers of accessions phenotyped across multiple environments and replications, which takes a significant amount of labor and resources. Machine learning (ML) methods are becoming more routine for phenotyping traits to save time and effort. This research aims to conduct GWAS on sudden death syndrome (SDS) of soybean [*Glycine max* L. (Merr.)]. This study uses disease severity from both visual field ratings and ML-based (using images) severity ratings collected from 473 accessions. Images were processed through an ML framework that identified soybean leaflets with SDS symptoms, and then disease severity was quantified on those leaflets into few classes. Both visual field ratings and image-based ratings identified significant single nucleotide polymorphism (SNP) markers associated with disease resistance. These significant SNP markers are either in the proximity of previously reported candidate genes for SDS, such as *ss715584164* and *ss715610404*, or near the potentially novel candidate genes, such as *ss715583703* and *ss715615734*. Within previously reported SDS quantitative trait loci there were significant SNPs from both visual rating and image-based ratings. The results of this study provide an exciting avenue for using ML to capture complex phenotypic traits from images to get comparable or more insightful results compared to subjective visual field stress phenotyping.