World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:17, No:06, 2023

Soybean Seed Composition Prediction From Standing Crops Using Planet Scope Satellite Imagery and Machine Learning

Authors : Supria Sarkar, Vasit Sagan, Sourav Bhadra, Meghnath Pokharel, Felix B.Fritschi

Abstract : Soybean and their derivatives are very important agricultural commodities around the world because of their wide applicability in human food, animal feed, biofuel, and industries. However, the significance of soybean production depends on the quality of the soybean seeds rather than the yield alone. Seed composition is widely dependent on plant physiological properties, aerobic and anaerobic environmental conditions, nutrient content, and plant phenological characteristics, which can be captured by high temporal resolution remote sensing datasets. Planet scope (PS) satellite images have high potential in sequential information of crop growth due to their frequent revisit throughout the world. In this study, we estimate soybean seed composition while the plants are in the field by utilizing PlanetScope (PS) satellite images and different machine learning algorithms. Several experimental fields were established with varying genotypes and different seed compositions were measured from the samples as ground truth data. The PS images were processed to extract 462 hand-crafted vegetative and textural features. Four machine learning algorithms, i.e., partial least squares (PLSR), random forest (RFR), gradient boosting machine (GBM), support vector machine (SVM), and two recurrent neural network architectures, i.e., long short-term memory (LSTM) and gated recurrent unit (GRU) were used in this study to predict oil, protein, sucrose, ash, starch, and fiber of soybean seed samples. The GRU and LSTM architectures had two separate branches, one for vegetative features and the other for textures features, which were later concatenated together to predict seed composition. The results show that sucrose, ash, protein, and oil yielded comparable prediction results. Machine learning algorithms that best predicted the six seed composition traits differed. GRU worked well for oil (R-Squared: of 0.53) and protein (R-Squared: 0.36), whereas SVR and PLSR showed the best result for sucrose (R-Squared: 0.74) and ash (R-Squared: 0.60), respectively. Although, the RFR and GBM provided comparable performance, the models tended to extremely overfit. Among the features, vegetative features were found as the most important variables compared to texture features. It is suggested to utilize many vegetation indices for machine learning training and select the best ones by using feature selection methods. Overall, the study reveals the feasibility and efficiency of PS images and machine learning for plot-level seed composition estimation. However, special care should be given while designing the plot size in the experiments to avoid mixed pixel issues.

Keywords : agriculture, computer vision, data science, geospatial technology
Conference Title : ICCVIP 2023 : International Conference on Computer Vision and Image Processing
Conference Location : Istanbul, Türkiye
Conference Dates : June 22-23, 2023