Detecting Crop Disease Before the Eye Can See It

Format: CS Scholarship Essay | Pattern: Research into Impact | ~500 words

My independent research project began with a simple question: could a convolutional neural network trained on satellite imagery reliably detect the early signs of crop disease before they were visible to field surveyors? The model I built, trained on publicly available Sentinel 2 data and a labeled dataset from an agricultural university, achieved 87 percent accuracy on a held-out validation set, roughly equivalent to the performance of an experienced agronomist reviewing the same imagery, at a fraction of the cost.

The project began as an extension of a geography assignment. I had been reading about satellite-based crop monitoring and was struck by a specific claim in a 2021 paper: that fungal infections in wheat fields show a measurable spectral signature in near-infrared imagery approximately 10 to 14 days before the infection becomes visible to the naked eye. If that signature could be reliably detected, farmers could apply targeted treatment before the infection spread to adjacent rows, potentially reducing both crop loss and pesticide use.

I spent four months acquiring and preprocessing Sentinel 2 imagery for a region of agricultural land in Kansas that had documented infection events in 2019 and 2020. The preprocessing involved atmospheric correction, band extraction, and the calculation of several vegetation indices that the literature suggested would be sensitive to early infection. I then trained a convolutional neural network on labeled examples of infected and uninfected field segments, using the 2019 data for training and the 2020 data for validation.

The 87 percent accuracy figure is promising but requires significant qualification. The validation set was drawn from the same geographic region and the same satellite instrument as the training set. Performance on imagery from different regions, different crop varieties, or different satellite instruments is unknown. A model that detects infection reliably in Kansas wheat may perform poorly on Iowa corn or Brazilian soybeans, and the practical utility of a crop disease detection system depends on precisely that kind of geographic and agronomic generalization.

The scholarship I am applying for would support my graduate studies in machine learning and remote sensing, where I intend to extend this work by acquiring training data from multiple geographic regions and crop types and evaluating the generalization problem systematically. The agricultural applications of satellite-based ML are substantial: early disease detection, irrigation optimization, yield prediction, and carbon sequestration monitoring are all problems where remotely sensed data combined with capable models could produce significant economic and environmental

benefits. I am at the beginning of that work, and I am asking for support to continue it.