Utilizing Hyperspectral Field Imagery for Accurate Southern Leaf Blight Severity Grading in Corn

Grace Vincent[1,2], Michael Kudenov[1,2], Peter Balint-Kurti[3], Ralph Dean[3], Cranos Williams[1,2,4]

[1] North Carolina Plant Sciences Initiative

[2] Department of Electrical and Computer Engineering, North Carolina State University, North Carolina, United States

[3] Department of Entomology and Plant Pathology, North Carolina State University, North Carolina, United States

[4] Department of Plant and Microbial Biology, North Carolina State University, North Carolina, United States

A paper from the Proceedings of the
16th International Conference on Precision Agriculture
21-24 July 2024
Manhattan, Kansas, United States

Abstract.

Crop disease detection using traditional scouting and visual inspection approaches can be laborious and time-consuming. Timely detection of disease and its severity over large spatial regions is critical for minimizing significant yield losses. Hyperspectral imagery has been demonstrated as a useful tool for a broad assessment of crop health. The use of spectral bands from hyperspectral data to predict disease severity and progression has been shown to have the capability of enhancing early disease detection, even before the onset of visible symptoms. In this study, off-axis hyperspectral imagery from entire fields are coupled with breeder-assigned plot-wise severity scores from a single growing season, to develop a machine-learning framework to detect and grade the severity of southern leaf blight infection in corn. The approach aimed to enhance the predictability of disease severity while assessing the interpretability of specific spectral wavelengths in relation to known biochemical processes. By employing unsupervised clustering techniques to isolate pixels associated with corn crops, a significant correlation between disease assessment and the spectra from segmented pixels was found. Feature reduction methods, e.g., Linear Discriminant Analysis (LDA), L1-Regularization, and Sequential Feature Selector, were implemented with a focus on identifying the most discriminative and influential wavelengths. These selected and extracted wavelengths were then evaluated using linear and non-linear regression-based models and quantified their effectiveness in identifying and grading disease in corn. The results showed that transforming the feature space with LDA achieves an R^2 value of 0.847. However, sacrificing some predictive power (R^2 of 0.607) enables the selection of wavelengths through Sequential Feature Search. Feature reduction identified spectral content in the 500-600 nm range through LDA loadings analysis and SFS evaluation, correlating with photosynthetic production indicators like

carotenoids and chlorophyll b, thereby highlighting the underlying biochemical mechanisms. The results show that hyperspectral imagery offers the unique advantage of scrutinizing the relative reflectance of specific wavelengths, enabling the capture of variations in disease grades and progression. These results hold significant promise for improving crop disease management, ultimately reducing crop losses and bolstering agricultural sustainability. Furthermore, the development of hyperspectral field imagery and machine learning models hold the potential for broader applications in monitoring and mitigating stressors in various crops, thereby advancing food security and promoting sustainability in agriculture.

Keywords.
Hyperspectral imaging, disease severity estimation, Southern Corn Leaf Blight, machine learning.

Introduction

Non-destructive imaging allows farmers and breeders to identify disease presence accurately and efficiently in crops and plants in a high throughput manner. Farmers and breeders have historically relied on scout teams to survey fields for disease presence within their crops demanding an enormous amount of time, labor, and expertise in distinguishing the variations in different stress responses. As agriculture is a multi-trillion dollar global industry and global crop losses due to biotic stressors are $60 billion, pointed and timely mitigation efforts are imperative to reducing any yield loss (Oerke, 2006; Reddy et al., 2009). Corn, a top five globally produced crop, faces the threat of 65 known pathogens causing an estimated 22.5% yield loss (Savary et al., 2019). Despite the variability in growth parameters and corn genotype, Southern leaf blight (SLB), a foliar disease, remains a significant concern due to its rapid development and potential for widespread transmission, as it is the 7th most destructive disease in the Southern United States (Mueller et al., 2020). Traditional non-destructive imaging systems (RGB imaging), are limited to assessing visible symptoms of stress response, which reduces their ability to recognize non-visual changes for a rapidly developing disease like SLB. Integrating imaging spectroscopy and near-infrared reflectance into disease recognition provides an objective measure, complementing traditionally subjective methods of manual field surveying and aiding in the deployment of efficient and accurate models for targeted mitigation efforts.

Notable advancements in automated leaf-level identification of diseases through machine learning have been facilitated by the accessibility of open-source comprehensive datasets like the PlantVillage dataset (Hughes and Salathé, 2015). Despite the strides made in leaf-level disease identification, models trained with images from controlled environments perform poorly in real-world settings, as shown by Mishra et al.'s 10% performance drop in field deployment (Mishra et al., 2020). Manual data collection through scout teams or UAV-based observations introduces inefficiencies, especially in capturing early-stage disease development. The limitations of overhead (nadir) observations with UAVs, as noted by Jia et al., hinder the effective capture of early-stage foliar diseases under the upper canopy (Jia et al., 2023; Zhang et al., 2019). This can impede timely mitigation efforts as foliar diseases tend to propagate from the basal leaves to the developing ear and finally to the flag lead of the plant (Ali et al., 2011). While both RGB and hyperspectral imaging are effective for recognizing disease (Mahlein et al., 2013), hyperspectral imaging provides insights into changes in pigmentation and internal structure (Gold et al., 2020), which is critical for asymptomatic recognition and assessing disease severity. Dense spectral information is often reduced into a transformation or a subset of influential wavelengths through methods such as computing spectral Vegetation Indices (VIs) (Jia et al., 2023; Khan et al., 2021), developing specific spectral Disease Indices (DIs) (Mahlein et al., 2013; Meng et al., 2020), and extracting key features (e.g., PCA) (Ravikanth et al., 2017). This refined data aids in assessing disease severity, offering deeper insights into the extent of damage, and informing mitigation strategies. While severity classification involves discretizing the levels of disease severity into distinct bins, whether this categorization is based on the development stage or the proportion of disease incident on plant leaves, estimating disease severity over a continuum of severity scores, provides a finer resolution and more nuanced understanding of the disease's impact and the

plant's response. For instance, UAV multispectral imaging effectively estimated the severity of potato late blight, with the NIR region showing significant differences for severity scores and achieved an R^2 value of 0.74 (Duarte-Carvajalino et al., 2018). Similarly, controlled hyperspectral imagery and computed VIs assessed the severity of Wheat Powdery Mildew, achieving a coefficient of determination (R^2) of 0.722, demonstrating the effectiveness of hyperspectral imagery in precise disease severity estimation (Khan et al., 2021). Severity estimation allows differentiation between asymptomatic stages, pre-visual symptoms, and early symptom development, particularly when leveraging hyperspectral information. This level of detail can provide valuable insights into disease progression and aid in decision-making for growers and breeders.

As such, this research explores the efficacy of off-axis polarized hyperspectral imaging in objectively identifying and quantifying the severity of Southern Corn Leaf Blight (SLB) within whole-field images. Departing from traditional disease detection and classification models, we employ a regression-based algorithms capable of distinguishing among a continuum of severity scores, thus providing detailed insights into the disease's progression. In hyperspectral imaging for plant phenotyping, challenges such as high dimensionality and a limited sample pool, contribute to the curse of dimensionality. To mitigate this issue, we implement non-overlapping subsampling of the plots, thereby augmenting the dataset. We examine various feature reduction techniques to enhance the prediction and estimation of severity and improve the interpretability of plant responses to SLB development. Results show high performance in the severity estimation of Southern Corn Leaf Blight, providing an efficient and effective assessment of disease at the plot level. Further, the reduced feature space is examined to identify wavelengths that are representative of Southern Leaf Blight development and connect these wavelengths to known underlying bio-physical processes. This has the potential to allow breeders and growers to implement targeted mitigation efforts by leveraging visible and nonvisible responses to disease incidence.

In summary, the main contributions of this research include:

- Assessing the utility of various hyperspectral image feature reduction techniques for estimating per-plot disease severity.
- Deriving an understanding of the underlying biological mechanisms driving the disease response.

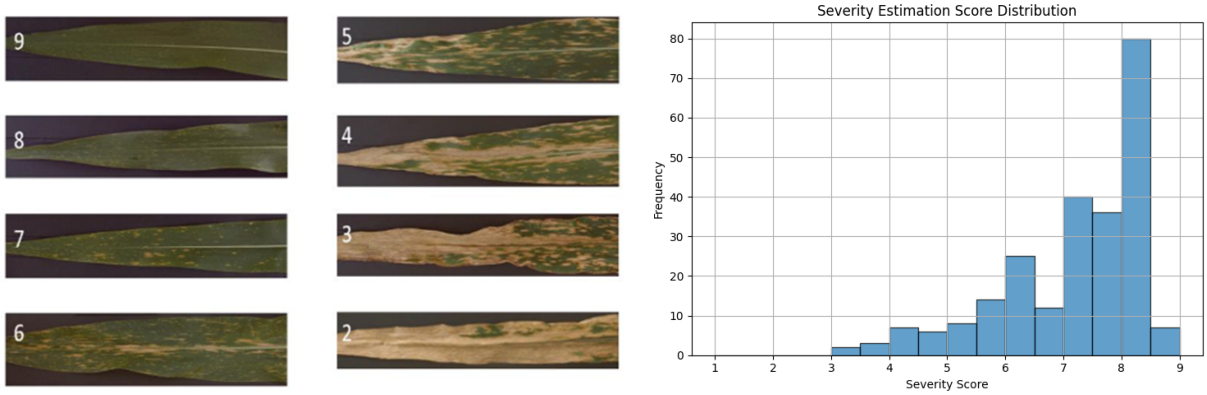
The rest of the paper is organized as follows. Methods presents the proposed pipeline of hyperspectral field images including preprocessing, unsupervised segmentation, and dimension reduction techniques used in this study. Finally, Results presents the results of the contributions listed above in detail.

Methods

Data

This work utilizes whole-field imagery from a hyperspectral imaging polarimeter mast camera (Kudenov et al., 2022) of corn development through the 2019 growing season. The plot fields were set up in two replications of the corn opposite each other. The hyperspectral imager captures one full replication in each image. Within each replication, 40 plots are identified with different corn genotypes with varying susceptibility to SLB. During the growing season, an expert breeder assigned corresponding plot-wise disease severity scores based on visual inspection following the procedure presented in (Sermons and Balint-Kurti, 2018). The scores are representative of the whole plot and assess the lesions on the ear leaf and the upper leaf. This work inspects images captured over three days with corresponding plot-wise scores. (a) Ear-leaf sample images (b) Severity score distributions

Figure 1a is a visual depiction of the varying scores that the breeder assigns.



(a) Ear-leaf sample images **(b) Severity score distributions**

Figure 1: (a)Ear-leaf samples of varying Southern Leaf Blight severity levels and (b) the severity score data distribution. The scores range from 9 to 1, the lower the more severe the symptom development.

The hyperspectral images provide spectral information from 150 evenly spaced spectral bands from 485 nanometers to 780 nanometers, capturing the visible light spectrum and near-infrared information. Hyperspectral information is utilized as it can capture the key vegetation characteristics based on the reflectance including the green peak (500-600 nm), chlorophyll well (600-680 nm), the red edge (680-750 nm), and the near-infrared (NIR) plateau (750-900 nm) (Teke et al., 2013). Within the visible spectrum, 400 nm to 700 nm, the leaf pigment and chlorophyll production dominates the response while the NIR spectrum, 700 nm to 1300 nm, provides insight into the internal structure of the leaf (Das et al., 2018).

In total, there are 6 hyperspectral datacubes, each containing the 40 designated plots, resulting in 240 plots used for analysis. The majority of plots exhibit minimal symptom development (i.e., a score >8), with scores ranging from 8.5 to 3, as depicted in Figure 1b. The dataset is stratified into training and testing sets based on a split of 80-20% according to severity scores. Further, the training set undergoes a stratified 5-fold cross-validation for model training.

Preprocessing

Figure 2: Sample image of the maize field in grayscale.

To account for the challenges regarding changes in illumination and incident angle of the sun, images were selected around noon. The hyperspectral images were then normalized with band min-max normalization technique (Cao et al., 2017; Suzuki et al., 2008) to account for varying illumination across the field or across time. The wavelengths below 500 nm and greater than 750 nm are omitted to reduce bias or artifacts at the end of the signatures, leaving 127 evenly spaced spectral bands. The spectral signatures are then smoothed using the widely adopted Savitzky-Golay filter (Bohnenkamp et al., 2019; Meng et al., 2020; Moghadam et al., 2017). This filter effectively reduces noise while preserving spectral information (Ravikanth et al., 2017).

Plot Identification

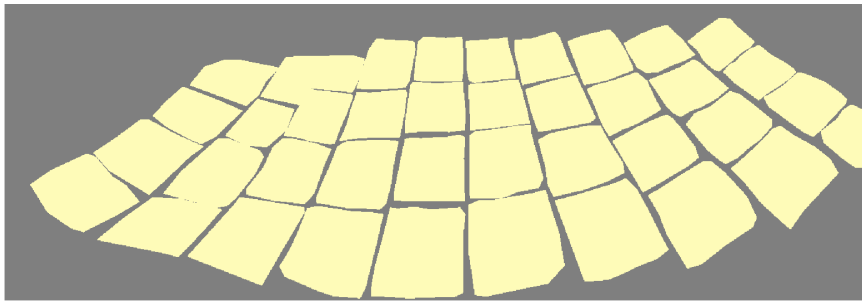


Figure 3: Sample field image with each plot colored. Each field image contains 40 plots.

Further preprocessing is required to identify the plots within the whole-field images (Figure 2) to pair to the plot-wise severity scores. Manual instance segmentation was employed to define the boundaries of the 40 individual plots within each image and match them to their individual severity scores. As a result, Figure 3 illustrates the delineated plot boundaries. This step was essential for subsequent analysis and for identifying trends within the plots that corresponded to the different severity scores. These plots will further be subset into a 3-by-3 grid to increase the number of plot images present within the dataset.

Crop Identification

Within the images are representations of corn crops and background objects such as soil and grass. The segmentation of crops from background objects aims to enhance the detection and description of SLB disease severity. Initial pixel masking employs computed Normalized-Difference Vegetation Index (NDVI) values to cluster vegetation-based pixels and identify impervious areas or image artifacts (Bah et al., 2018; Lassiter and Darbari, 2020; Lu et al., 2022; Suzuki et al., 2008). NDVI is derived from the normalized difference between the red and NIR bands and indicates vegetation density or 'greenness' within pixels or regions (Carlson and Ripley, 1997). While NDVI thresholding commonly distinguishes vegetation from background pixels, its effectiveness is primarily observed in settings with clear vegetation-background boundaries, such as urban environments or controlled imaging conditions. Evaluating unsupervised clustering with and without thresholding will assess its adaptability to complex environments lacking distinct boundaries. The limitations of single wavelengths spectral vegetation indices (Vis) in accurately classifying diseased or damaged vegetation underscore the need for refined segmentation techniques.

Following initial NDVI masking, unsupervised K-Means clustering (Sinaga and Yang, 2020) isolates corn crop pixels to generate a crop mask (Hamuda et al., 2016), highlighting the value of using both the NDVI mask. The optimal number of clusters is determined by balancing unsupervised metrics and visually assessing the resulting clusters. These metrics include the within-cluster sum of squares to minimize variance within a cluster, the Silhouette Score to maximize the similarity of data within clusters and dissimilarity between clusters, and the Bayes Information Criterion.

Feature Reduction

Hyperspectral information is inherently high-dimensional with a broad spectral range and small spectral resolutions, i.e., 500 nm - 750 nm with a resolution of ~2 nm. These spectral bands experience multicollinearity as there is a strong correlation between each band necessitating a reduced feature space (Chan et al., 2022). (a) Ear-leaf sample images (b) Severity score distributions

Figure 1

Feature Selection

Using feature selection remains a popular method for performing dimension reduction for hyperspectral imagery. Feature selection is a method of feature reduction that subsets the original set of features maintaining the physical explanations of the features (Vijouyeh and Taskin, 2016).

Feature selection contextualizes the reduced feature space, however, it also produces feature subsets that are overly specific to a given task (Brown et al., 2012).

This paper explores four techniques: correlation-based, mutual information (Brown et al., 2012), Sequential Feature Selector (SFS) (Kwak and Chong-Ho Choi, 2002), and l-1 regularized linear regression (LASSO) (Li et al., 2021). Correlation-based feature selection identifies features that have the highest relationship with the target variable, i.e., disease severity score, by calculating the correlation coefficient. Mutual information calculates the amount of information that the features provide about the disease severity score. SFS performs forward and backward feature selection by adding and removing features from the selected subset in a greedy manner. SFS seeks to minimize the criterion over all the achievable feature subsets. The final feature selection method is LASSO, it is a common technique as it imposes l-1 norm constraints on regression coefficients to perform both regression and feature selection simultaneously. This work utilizes the feature selection ability of LASSO to identify the features with the largest absolute coefficients.

Feature Extraction

In contrast, feature extraction is a transformation of the original feature space to make the classes more separable by extracting useful properties of the features (Vijouyeh and Taskin, 2016). This method of transforming the feature space removes contextual information from the reduced space but often reduces analogous features increasing the separability/discrimination between features (Zhao and Du, 2016). Principal component analysis (PCA) captures the most significant variations in the data without consideration for class associations. While there exist unsupervised and semi-supervised methods of feature extraction, this work focuses on utilizing supervised techniques, i.e., Linear Discriminant Analysis (LDA) (Fang et al., 2014). LDA is a common feature reduction method that maximizes the ratio of between-class variance to within-class variance as such the loadings in LDA are the eigenvectors of the between-class scatter matrix multiplied by the inverse of the within-class scatter matrix.

Evaluation

To assess the various feature reduction techniques, both Linear Regression (Seber and Lee, 2003) and Support Vector Regression (SVR) are used to handle the task of scoring the severity of the disease. These models investigate linear and non-linear relationships between the severity scores and the hyperspectral reflectance. The Ezekiel adjusted R^2 metric (Ezekiel, 1929) and a Root Mean Squared Error (RMSE) metric are utilized to measure the effectiveness of the techniques. These metrics are chosen because of their ability to account for the fluctuations in the number of features, i.e., spectral bands, (F) and samples, i.e., subplots, (S) employed during the evaluation process, providing a more robust and fair assessment, as seen in Equation (1).

$$\text{Adj } R^2 = 1 - \frac{S-1}{S-F-1} \times \left(1 - R^2\right) \quad (1)$$

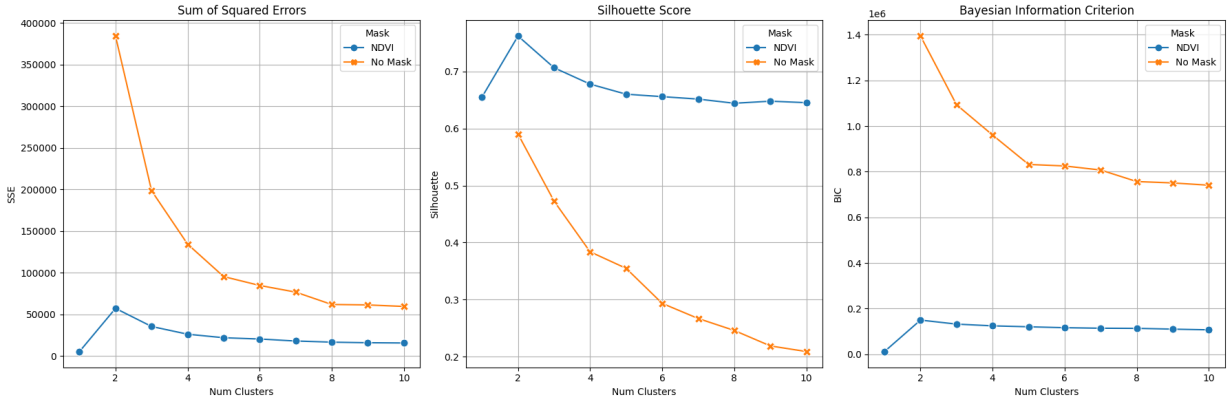
To determine the optimal feature reduction technique, model, and number of features a 5-fold cross-validation is performed on 80% of the total subplots. The Adjusted R^2 and RMSE metrics are averaged over the 5-folds to find the top-performing features for estimating severity and then tested on the remaining 20% of subplots for final performance metrics.

Results

NDVI masking distinguishes crops from background in complex environments.

Unsupervised K-Means clustering was leveraged to distinguish the crops from all other background elements in the images. Additionally, an initial masking of NDVI values was implemented to enhance the discrimination of the K-Means clustering and the applicability of NDVI masking to complex images. Three different metrics, the Sum of Squared Errors (SSE), the Silhouette Score, and the Bayesian Information Criterion (BIC) are evaluated and plotted in Figure 4. Clustering an initial NDVI mask reduces the SSE and BIC while maximizing the Silhouette

Score, outperforming clustering without an initial mask. Traditionally, identifying the optimal number of clusters in unsupervised clustering requires identifying the 'elbow point' in the SSE plot, where adding more clusters no longer significantly reduces the SSE. Based on the elbow method, four clusters would be optima; however, with just one cluster (the NDVI mask) achieves the lowest SSE value (Figure 4a). A similar pattern is observed with the BIC, indicating that the initial NDVI mask minimizes separation within clusters. Yet, just two clusters the Silhouette Score is maximized (Figure 4b).

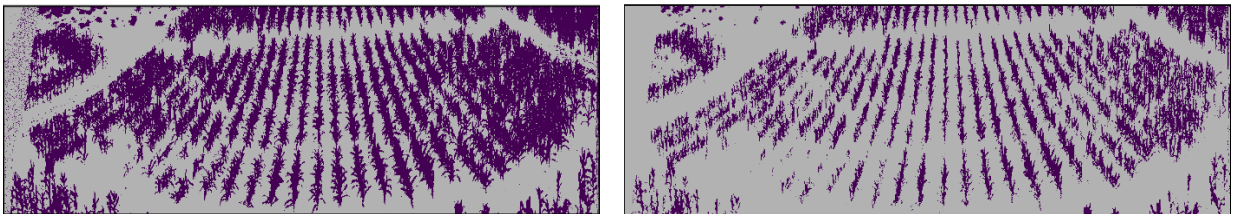


(a) Sum of Squared Errors **(b) Silhouette Score** **(c) Bayesian Information Criterion**

Figure 4: Resulting unsupervised clustering metrics for both clustering after an initial NDVI mask (blue) and without a mask (orange). The initial NDVI mask metrics are indicated by a Number of Clusters = 1.

Using visual inspection, it has been found that the initial NDVI mask can distinguish the crop boundaries from background elements better than an additional two clusters (Figure 5). The visual inspection also shows that two clusters have incomplete segmentation masks, meaning that there are pixels within the crop that are expected to be labeled as crop but are classified as background. This is less evident with the NDVI mask. Through the balance of unsupervised metrics and visual assessment, an NDVI mask is determined to be optimal for distinguishing between the corn crops and all other pixel classes ((a) Initial NDVI Mask (b) 2 Clusters

Figure 5a).

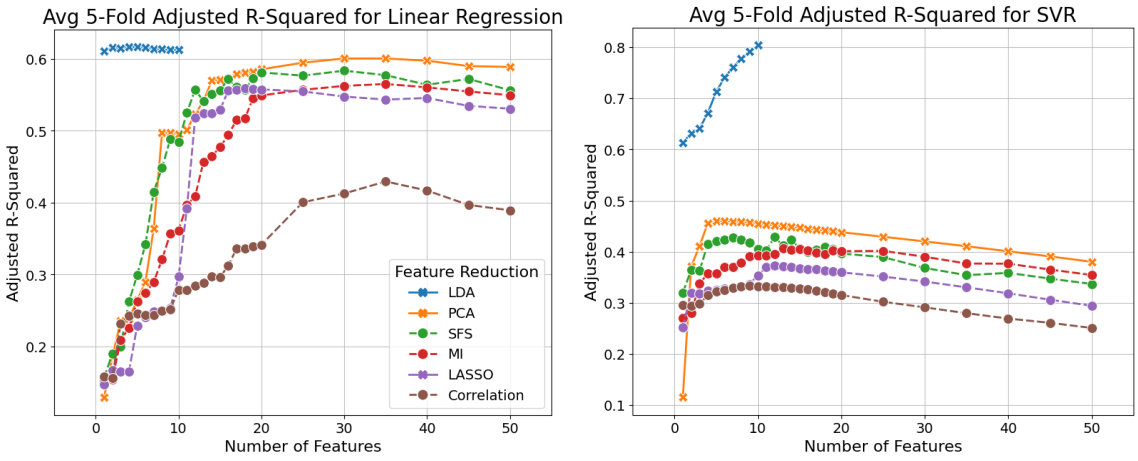


(a) Initial NDVI Mask **(b) 2 Clusters**

Figure 5: Visual comparison of (a) an NDVI mask and (b) 2 clusters after an initial NDVI mask. Purple corresponds to what is determined to be crops.

Cross-validation identifies optimal feature reduction methods.

Evaluating the 5-fold cross-validation results for both linear regression and Support Vector Regression (SVR) models, Linear Discriminant Analysis (LDA) emerges as the top-performing feature reduction technique, as illustrated in Figure 6. LDA demonstrated peak performance when 10 features were used in an SVR model, significantly outperforming all other reduction techniques (Figure 6b). Sequential Feature Search (SFS) yields an average adjusted R^2 value of ~0.55 in the 5-fold cross-validation, peaking at 30 features and an Adj. R^2 of 0.58 (Figure 6a). With fewer than 20 features, LASSO features achieved an average adjusted R^2 value greater than 0.5 with a linear regressor.



(a) Linear Regression **(b) Support Vector Regression**

Figure 6: Average Adjusted R2 for an increasing number of features for the 5-Fold Cross-Validation. Feature extraction methods are indicated with solid lines and X markers and feature selection methods are indicated with dashed lines and O markers.

Feature extraction outperforms selection methods in estimating disease severity

The top-performing combination of feature reduction methods, number of features, and model for the severity estimation task are then tested against the holdout test dataset. The results indicate that feature extraction techniques, specifically Linear Discriminant Analysis (LDA), outperform the top-performing feature selection techniques such as SFS or LASSO.

Table 1: Results of top-performing feature extraction (LDA) and feature selection (LASSO and SFS) techniques when evaluated on the holdout test set. The adjusted R^2, R^2, and RMSE metrics are reported for both a linear regressor and SVR.

Model	Reduction	# of Features	Adj. R^2	R^2	RMSE
Support Vector Regression	LDA	10	0.847	0.906	0.969
Linear Regression	SFS	30	0.607	0.632	1.553
Linear Regression	LASSO	18	0.550	0.589	1.660

Using LDA with 10 features and an SVR model achieved an adjusted R^2 0.847 and minimizes the RMSE to 0.969, indicating high predictive accuracy (Table 1). This highlights the effectiveness of LDA in capturing the most informative features that contribute significantly to the model's performance.

In contrast, applying the 30 features selected through SFS to a linear regression model resulted in an adjusted R^2 value of 0.607 and an RMSE of 1.553 (Table 1). With 12 fewer features, LASSO achieves an adjusted R^2 value of 0.550 on the holdout set. These feature selection methods (i.e., SFS and LASSO) are able to explain portions of the variance in the disease severity and contain some predictive power, but still fall short of the performance as a result of LDA extracted features.

Linking bio-physical processes to the reduced feature spaces.

Feature extraction methods are a transformation of the original feature space, but viewing the projection matrices provides insight into the spectral features that are important to SLB scores. By examining the 10 LDA components and calculating the mean and standard deviation of their loadings, significant spectral features are identified based on magnitudes greater than one standard deviation away from the loading's mean. These prevalent spectral features are plotted in Figure 7, where the height of the bars indicates how many loadings the spectral feature is influential in.

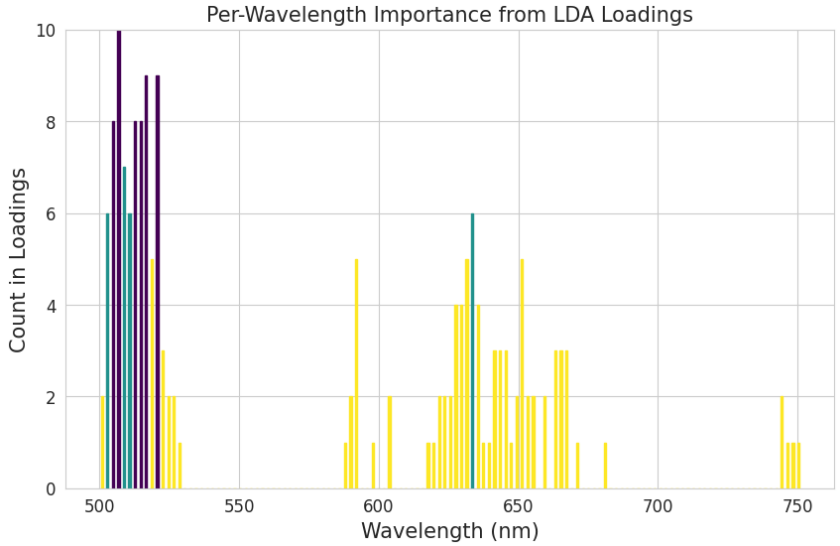


Figure 7: Identified prevalent and influential spectral features within the LDA loadings. Prevalent features are more than one standard deviation from the per-loading mean. The height of each spectral feature corresponds to the total number of loadings the spectral feature is prevalent in, maximum value is 10. Purple-colored bars represent prevalent wavelengths in 8 or more components, teal-colored bars represent prevalent in 6 or more components, and yellow is all other wavelengths.

From this analysis, it is observed that the most prevalent spectral features (greater than one standard deviation away from the per-loading mean for 8 loadings) range from 502 nm to 520 nm with a peak at 514 nm. These features are depicted by purple-colored bars, fall within the green peak of the visible spectrum, known for estimating chlorophyll content, particularly as this is the region of peak Chlorophyll b absorption (Teke et al., 2013). Additionally, spectral features prevalent in 6 or more loadings, shown in teal-colored bars, also highlight the range from 502 nm to 520 nm. Notably, there is a single prevalent spectral feature at 633 nm in 6 loadings. This region, spanning 625 nm to 675 nm, captures moderately influential spectral bands but it also within the broader chlorophyll well.

Despite a diminished predictive power, the top-performing feature selection technique, SFS, directly connects with the context of the original feature space; that is, the selected spectral features maintain their relation to any bio-physical processes. As such, the 30 selected wavelengths can be located, and regions or clusters of wavelengths that were deemed influential can be identified (Figure 8). Two regions stand out as particularly influential as there due to simple clustering of spectral bands: 500 nm – 598 nm and 682 nm – 742 nm. Notably, wavelengths 617 nm and 629 nm were excluded from a cluster due to their coefficients' magnitude being significantly smaller in scale than the other wavelengths.

The region 500 nm – 598 nm is dominated by Carotenoid and Anthocyanin content within the green peak (Gitelson et al., 2006). The cluster that ranges from 682 nm to 742 nm captures information within the chlorophyll well, which describes chlorophyll content, and the red-edge region is known for its ability to identify and grade vegetation stress as it captures changes within internal structure and chlorophyll content(Goswami et al., 2021; Wiefels and Baroja, 2022). Primarily these selected spectral features capture changes within the pigmentation of the leaf, but the region of the red edge and NIR plateau can begin to provide insights into internal changes.

There is a region of identified spectral features that consistently appeared influential across both the extracted and selected features. Specifically, the region corresponding to the green peak and Chlorophyll b absorption exhibited notable importance. This finding aligns with previous research on classifying the severity of Southern Leaf Blight (SLB) in corn using vegetation indices, where the Chlorophyll content was also identified as significant (Lv et al., 2023; Safir, 1972). Given that SLB disease symptoms primarily result from chlorophyll degradation, the results underscore the importance of this spectral signature in modeling disease progression.

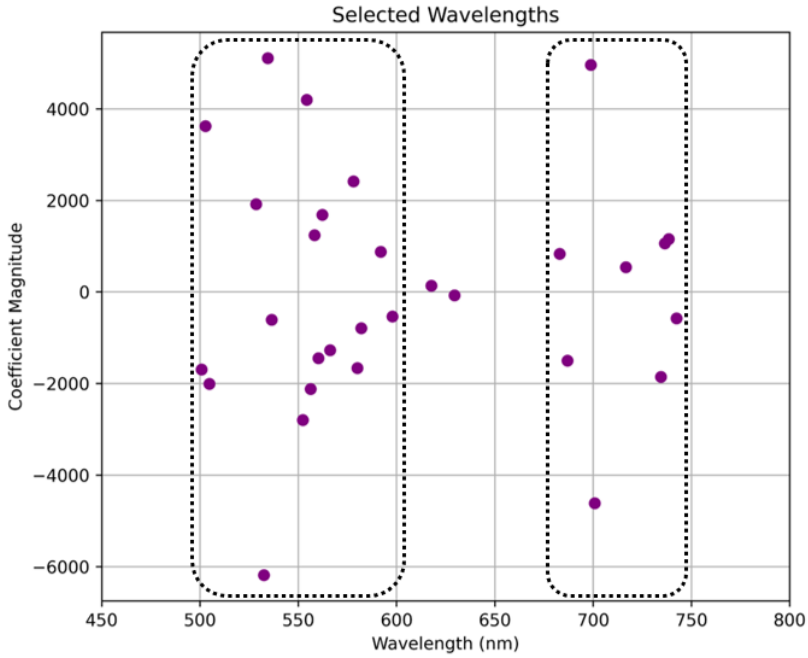


Figure 8: 30 selected wavelengths or features from Sequential Feature Search.

Conclusion

In this study, the aim was to design an automated pipeline for objectively assessing the severity of Southern Corn Leaf Blight (SLB) in high-throughput whole-field hyperspectral images. By evaluating spectral features indicative of SLB's presence and progression, the goal was to improve subsequent analysis for forecasting disease progression and deepen the understanding of the bio-physical responses to the disease. The findings reveal that while feature selection techniques offer direct connections to contextual interpretations, feature extraction methods demonstrate higher predictive power. Through the examination of the loadings of Linear Discriminant Analysis (LDA), insights were gained into the spectral features of importance, uncovering consistent features across both feature selection and extraction techniques. Notably, the region of the green peak (500 nm to 600 nm) highlights the hindrance of Chlorophyll b production as Southern Leaf Blight progresses. By analyzing the spectral features identified as influential and important, a balance is struck between interpretable features and accurate assessment of disease severity. This integrated approach paves the way for a more comprehensive understanding of SLB dynamics and facilitates the development of effective strategies for disease management and crop protection.

Acknowledgments

The authors want to acknowledge the Foundation for Food and Agriculture Research Fellowship and the Graduate Fellowship for STEM Diversity for their financial support. Special thanks to the EnBiSys Lab and the North Carolina Plant Sciences Initiative for access to the High-Performance Cluster, and to Dr. Michael Kudenov and Dr. Peter Balint-Kurti for their contributions in facilitating field experiments and providing essential datasets.

References

Ali, F., Muneer, M., Rahman, H., Noor, M., Yan, J., 2011. Heritability estimates for yield and related traits based on testcross progeny performance of resistant maize inbred lines.

Bah, M., Hafiane, A., Canals, R., 2018. Deep Learning with Unsupervised Data Labeling for Weed Detection in Line Crops in UAV Images. Remote Sens. 10, 1690. https://doi.org/10.3390/rs10111690

Bohnenkamp, D., Behmann, J., Mahlein, A.-K., 2019. In-Field Detection of Yellow Rust in Wheat on the Ground Canopy and UAV Scale. Remote Sens. 11, 2495. https://doi.org/10.3390/rs11212495

Brown, G., Pocock, A., Zhao, M.-J., Lujan, M., 2012. Conditional Likelihood Maximisation: A Unifying Framework for Information Theoretic Feature Selection.

Cao, F., Yang, Z., Ren, J., Jiang, M., Ling, W.-K., 2017. Does Normalization Methods Play a Role for Hyperspectral Image Classification?

Carlson, T.N., Ripley, D.A., 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sens. Environ. 62, 241–252. https://doi.org/10.1016/S0034-4257(97)00104-1

Chan, J.Y.-L., Leow, S.M.H., Bea, K.T., Cheng, W.K., Phoong, S.W., Hong, Z.-W., Chen, Y.-L., 2022. Mitigating the Multicollinearity Problem and Its Machine Learning Approach: A Review. Mathematics 10, 1283. https://doi.org/10.3390/math10081283

Das, B., Mahajan, G.R., Singh, R., 2018. Hyperspectral Remote Sensing: Use in Detecting Abiotic Stresses in Agriculture, in: Advances in Crop Environment Interaction. Springer Singapore, pp. 317–335. https://doi.org/10.1007/978-981-13-1861-0_12

Duarte-Carvajalino, J.M., Alzate, D.F., Ramirez, A.A., Santa-Sepulveda, J.D., Fajardo-Rojas, A.E., Soto-Suárez, M., 2018. Evaluating late blight severity in potato crops using unmanned aerial vehicles and machine learning algorithms. Remote Sens. 10. https://doi.org/10.3390/rs10101513

Ezekiel, M., 1929. The Application of the Theory of Error to Multiple and Curvilinear Correlation.

Fang, Y., Li, H., Ma, Y., Liang, K., Hu, Y., Zhang, S., Wang, H., 2014. Dimensionality Reduction of Hyperspectral Images Based on Robust Spatial Information Using Locally Linear Embedding. IEEE Geosci. Remote Sens. Lett. 11, 1712–1716. https://doi.org/10.1109/LGRS.2014.2306689

Gitelson, A.A., Keydan, G.P., Merzlyak, M.N., 2006. Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. Geophys. Res. Lett. 33, 2006GL026457. https://doi.org/10.1029/2006GL026457

Gold, K.M., Townsend, P.A., Chlus, A., Herrmann, I., Couture, J.J., Larson, E.R., Gevens, A.J., 2020. Hyperspectral Measurements Enable Pre-Symptomatic Detection and Differentiation of Contrasting Physiological Effects of Late Blight and Early Blight in Potato. Remote Sens. 12, 286. https://doi.org/10.3390/rs12020286

Goswami, J., Das, R., Sarma, K.K., Raju, P.L.N., 2021. Red Edge Position (REP), an Indicator for Crop Stress Detection: Implication on Rice (Oryza sativa L). Int. J. Environ. Clim. Change 88–96. https://doi.org/10.9734/ijecc/2021/v11i430396

Hamuda, E., Glavin, M., Jones, E., 2016. A survey of image processing techniques for plant extraction and segmentation in the field. Comput. Electron. Agric. 125, 184–199. https://doi.org/10.1016/j.compag.2016.04.024

Hughes, D.P., Salathé, M., 2015. An open access repository of images on plant health to enable the development of mobile disease diagnostics.

Jia, X., Yin, D., Bai, Yali, Yu, X., Song, Y., Cheng, M., Liu, S., Bai, Yi, Meng, L., Liu, Y., Liu, Q., Nan, F., Nie, C., Shi, L., Dong, P., Guo, W., Jin, X., 2023. Monitoring Maize Leaf Spot Disease Using Multi-Source UAV Imagery. Drones 7, 650. https://doi.org/10.3390/drones7110650

Khan, I.H., Liu, Haiyan, Li, W., Cao, A., Wang, X., Liu, Hongyan, Cheng, T., Tian, Y., Zhu, Y., Cao, W., Yao, X., 2021. Early Detection of Powdery Mildew Disease and Accurate Quantification of Its Severity Using Hyperspectral Images in Wheat. Remote Sens. 13, 3612. https://doi.org/10.3390/rs13183612

Kudenov, M.W., Altaqui, A., Williams, C., 2022. Practical spectral photography II: snapshot spectral imaging using linear retarders and microgrid polarization cameras. Opt. Express 30, 12337. https://doi.org/10.1364/OE.453538

Kwak, N., Chong-Ho Choi, 2002. Input feature selection for classification problems. IEEE Trans. Neural Netw. 13, 143–159. https://doi.org/10.1109/72.977291

Lassiter, A., Darbari, M., 2020. Assessing alternative methods for unsupervised segmentation of urban vegetation in very high-resolution multispectral aerial imagery. PLOS ONE 15, e0230856. https://doi.org/10.1371/journal.pone.0230856

Li, F., Lai, L., Cui, S., 2021. On the Adversarial Robustness of LASSO Based Feature Selection. IEEE Trans. Signal Process. 69, 5555–5567. https://doi.org/10.1109/TSP.2021.3115943

Lu, Z., Qi, L., Zhang, H., Wan, J., Zhou, J., 2022. Image Segmentation of UAV Fruit Tree Canopy in a Natural Illumination Environment. Agriculture 12, 1039. https://doi.org/10.3390/agriculture12071039

Lv, Z., Meng, R., Chen, G., Zhao, F., Xu, B., Zhao, Y., Huang, Z., Zhou, L., Zeng, L., Yan, J., 2023. Combining multiple spectral enhancement features for improving spectroscopic asymptomatic detection and symptomatic severity classification of southern corn leaf blight. Precis. Agric. 24, 1593–1618. https://doi.org/10.1007/s11119-023-10010-2

Mahlein, A.-K., Rumpf, T., Welke, P., Dehne, H.-W., Plümer, L., Steiner, U., Oerke, E.-C., 2013. Development of spectral indices for detecting and identifying plant diseases. Remote Sens. Environ. 128, 21–30. https://doi.org/10.1016/j.rse.2012.09.019

Meng, R., Lv, Z., Yan, J., Chen, G., Zhao, F., Zeng, L., Xu, B., 2020. Development of Spectral Disease Indices for Southern Corn Rust Detection and Severity Classification. Remote Sens. 12, 3233. https://doi.org/10.3390/rs12193233

Mishra, S., Sachan, R., Rajpal, D., 2020. Deep Convolutional Neural Network based Detection System for Real-time Corn Plant Disease Recognition. Procedia Comput. Sci. 167, 2003–2010. https://doi.org/10.1016/j.procs.2020.03.236

Moghadam, P., Ward, D., Goan, E., Jayawardena, S., Sikka, P., Hernandez, E., 2017. Plant Disease Detection Using Hyperspectral Imaging, in: 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA). Presented at the 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), IEEE, Sydney, NSW, pp. 1–8. https://doi.org/10.1109/DICTA.2017.8227476

Mueller, D.S., Wise, K.A., Sisson, A.J., Allen, T.W., Bergstrom, G.C., Bissonnette, K.M., Bradley, C.A., Byamukama, E., Chilvers, M.I., Collins, A.A., Esker, P.D., Faske, T.R., Friskop, A.J., Hagan, A.K., Heiniger, R.W., Hollier, C.A., Isakeit, T., Jackson-Ziems, T.A., Jardine, D.J., Kelly, H.M., Kleczewski, N.M., Koehler, A.M., Koenning, S.R., Malvick, D.K., Mehl, H.L., Meyer, R.F., Paul, P.A., Peltier, A.J., Price, P.P., Robertson, A.E., Roth, G.W., Sikora, E.J., Smith, D.L., Tande, C.A., Telenko, D.E.P., Tenuta, A.U., Thiessen, L.D., Wiebold, W.J., 2020. Corn Yield Loss Estimates Due to Diseases in the United States and Ontario, Canada, from 2016 to 2019. Plant Health Prog. 21, 238–247. https://doi.org/10.1094/PHP-05-20-0038-RS

Oerke, E.-C., 2006. Crop losses to pests. J. Agric. Sci. 144, 31–43. https://doi.org/10.1017/S0021859605005708

Ravikanth, L., Jayas, D.S., White, N.D.G., Fields, P.G., Sun, D.-W., 2017. Extraction of Spectral Information from Hyperspectral Data and Application of Hyperspectral Imaging for Food and Agricultural Products. Food Bioprocess Technol. 10, 1–33. https://doi.org/10.1007/s11947-016-1817-8

Reddy, D.V.R., Sudarshana, M.R., Fuchs, M., Rao, N.C., Thottappilly, G., 2009. Genetically Engineered Virus-Resistant Plants in Developing Countries: Current Status and Future Prospects, in: Advances in Virus Research. Elsevier, pp. 185–220. https://doi.org/10.1016/S0065-3527(09)07506-X

Safir, G.R., 1972. Spectral Reflectance and Transmittance of Corn Leaves Infected with Helminthosporium maydis. Phytopathology 62, 1210. https://doi.org/10.1094/Phyto-62-1210

Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., Nelson, A., 2019. The global burden of pathogens and pests on major food crops. Nat. Ecol. Evol. 3, 430–439. https://doi.org/10.1038/s41559-018-0793-y

Seber, G.A.F., Lee, A.J., 2003. Linear Regression Analysis, 1st ed, Wiley Series in Probability and Statistics. Wiley. https://doi.org/10.1002/9780471722199

Sermons, S., Balint-Kurti, P., 2018. Large Scale Field Inoculation and Scoring of Maize Southern Leaf Blight and Other Maize Foliar Fungal Diseases. BIO-Protoc. 8. https://doi.org/10.21769/BioProtoc.2745

Sinaga, K.P., Yang, M.-S., 2020. Unsupervised K-Means Clustering Algorithm. IEEE Access 8, 80716–80727. https://doi.org/10.1109/ACCESS.2020.2988796

Suzuki, Y., OKAMOTO, H., KATAOKA, T., 2008. Image Segmentation between Crop and Weed using Hyperspectral Imaging for Weed Detection in Soybean Field. Environ. Control Biol. 46, 163–173. https://doi.org/10.2525/ecb.46.163

Teke, M., Deveci, H.S., Haliloglu, O., Gurbuz, S.Z., Sakarya, U., 2013. A short survey of hyperspectral remote sensing applications in agriculture, in: 2013 6th International Conference on Recent Advances in Space Technologies (RAST). IEEE, pp. 171–176. https://doi.org/10.1109/RAST.2013.6581194

Vijouyeh, H.G., Taskin, G., 2016. A comprehensive evaluation of feature selection algorithms in hyperspectral image classification, in: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Presented at the IGARSS 2016 - 2016 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Beijing, China, pp. 489–492. https://doi.org/10.1109/IGARSS.2016.7729121

Wiefels, A., Baroja, C., 2022. Red Edge Detects Vegetative Stress Earlier in Plant Growth Cycle [WWW Document]. MundoGEO. URL https://mundogeo.com/en/2022/02/09/red-edge-detects-vegetative-stress-earlier-in-plant-growth-cycle/ (accessed 6.5.24).

Zhang, X., Han, Liangxiu, Dong, Y., Shi, Y., Huang, W., Han, Lianghao, González-Moreno, P., Ma, H., Ye, H., Sobeih, T., 2019. A deep learning-based approach for automated yellow rust disease detection from high-resolution hyperspectral UAV images. Remote Sens. 11. https://doi.org/10.3390/rs11131554

Zhao, W., Du, S., 2016. Spectral–Spatial Feature Extraction for Hyperspectral Image Classification: A Dimension Reduction and Deep Learning Approach. IEEE Trans. Geosci. Remote Sens. 54, 4544–4554. https://doi.org/10.1109/TGRS.2016.2543748