Modeling Diseases in Corn Leaves Using Computer Vision

Sam Borremans*[1], Samuel Orellana Mateo*[1], Yash Singam*[1], Samir Travers*[1], Benjamin Yan*[1]

June 17, 2024

Abstract

This project explores the dual challenge of developing effective convolutional neural networks (CNNs) for disease detection in corn leaves and addressing critical issues in dataset quality. Utilizing the Maize Leaf Disease dataset, which contains around 4000 images categorized into four classes—three diseased and one healthy—we initially found that the models were prone to overfitting due to category-dependent backgrounds unique to each category. This paper details our efforts to design and compare CNN models, including ResNet, VGG16, EfficientNet, MobileNet, DenseNet, LeNet, and ShuffleNet, while simultaneously tackling the dataset's background uniformity problem. By employing color-based segmentation to standardize image backgrounds, we ensured that the models focused on the leaf features themselves. Our results demonstrate significant improvements in model accuracy and reliability, highlighting the importance of robust dataset preprocessing alongside advanced model architecture design for effective disease detection.

---

* Equal Contribution.

[1] Undergraduate student at Duke University.

Sam Borremans
sam.borremans@duke.edu

Samuel Orellana Mateo
samuel.orellanamateo@duke.edu

Yash Singam
yash.singam@duke.edu

Samir Travers
samir.travers@duke.edu

Benjamin Yan
benjamin.yan@duke.edu

1. Introduction

In 2022, the United States experienced a significant loss of around 380 million bushels of corn [1] due to various diseases. Specifically, 33.6 million bushels were lost to three aboveground diseases: common rust, gray leaf spot, and blight. These diseases can be identified by examining the leaves of affected plants, a relatively straightforward process. However, implementing field-wide coverage for detection, especially considering that corn fields averaged 725 acres in 2017 [2], can be a slow and labor-intensive process. Therefore, we sought to build models utilizing computer vision technology to swiftly identify and flag unhealthy corn affected by the diseases mentioned above.

The three diseases were chosen because they are aboveground diseases and are particularly prevalent in the United States. The table below summarizes their effects, location, how they are detected, and their impact on quality:

	Common Rust [3][4]	**Gray Leaf Spot** [5][6][7]	**Blight** [8][9]
Effects	Under severe conditions, can lead to leaf chlorosis (yellowing) and early death.	Reduces the photosynthetic area on leaves. Lesions may coalesce and kill leaves (grayish appearance due to fungal spores).	Reduces the photosynthetic area on leaves. Lesions produce olive-green or black fungal spores (high humidity), visible with a hand lens
Location	South U.S. and Mexico. Spreads to the north in summer through wind-carried spores. Found when the weather is cool (limited at 80◦ F) and moist.	Primarily U.S. corn belt. Found in locations with wet and humid weather. Usually, the outbreaks are not widespread.	Midwestern U.S. Found in locations with moderate temperatures (64 – 80 °F), wet and humid weather, or long periods of moisture.
Detection	Small and either round or elongated brown pustules form on both leaf surfaces and other aboveground parts.	Starts as small necrotic spots with halos on corn leaves and expands into rectangular gray-to-brown lesions with parallel edges.	Canoe-shaped lesions, 1-6 inches long, initially bordered by gray-green margins before becoming tan-colored. Starts on lower leaves and spreads to upper ones.
Impact on quality	May not have a direct impact in terms of grain characteristics, but can weaken plants and reduce overall yield.	Can significantly affect yields, especially in susceptible hybrids, and can lead to rotting and lodging.	Reduction in quality in sweet and silage corn, but minimal yield losses. Can become significant with susceptible hybrids or inbreds.

Figure 1: *Descriptions of the common rust, gray leaf spot, and blight diseases*

2. Dataset

The data utilized in this study was sourced from the Corn or Maize Leaf Disease dataset [10], which encompassed four distinct classes of corn leaves. One class represented healthy leaves, while the other classes consisted of leaves afflicted by each of the described diseases. The distribution of these categories is illustrated on the right.

All images were manually captured and exhibit variations in zoom level, dimensions, and orientation (both horizontal and vertical). Moreover, the backgrounds of the images vary; some feature other corn leaves, some have a gray background (typical in lab-controlled images), some exhibit a blurred background, and others lack a background altogether (black pixels). Consequently, a good model must be able to delineate the edges of the leaf in each image.

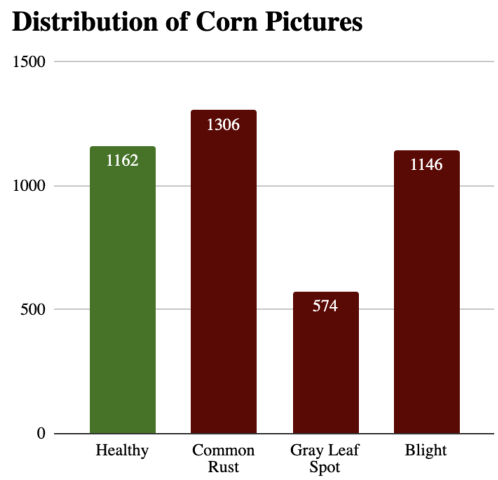


Figure 2: *Distribution of the dataset*

3. Models

3.1. Model Types

To detect and analyze corn leaves, our models employ neural networks, a common approach in computer vision. We opted for convolutional neural networks implemented using PyTorch [11], and used the following models, based on existing literature exploring similar use cases:

Model Name	Description of Literature
ResNet [12]	Godi et al. utilized a ResNet model (ResNet 152 v2) to predict plant leaf diseases, achieving 95% accuracy on a dataset of 7000 images, surpassing previous models.
VGG16 [13]	Bhatt et al. used a VGG16 model to classify corn diseases in 2000 images and compare its performance to that of Inception-v2, ResNet50, and MobileNet-v1
EfficientNet [14]	Sun et al. leveraged FL-EfficientNet CNN architecture to classify corn diseases and compare its performance to that of ResNet50, DenseNet169
LeNet [15]	Wei et al. employed a LeNet-5 architecture to identify gases for electronic noses, achieving a final gas identification accuracy rate of 98.67%.
ShuffleNet [16]	Zhang et al. created an instance of the ShuffleNet CNN architecture optimized for mobile devices that use pointwise group convolution and channel shuffle.
DenseNet [17]	Gao et al. designed a DenseNet architecture utilizing a feed-forward design with connections between all the layers to achieve high accuracy.
MobileNet [18]	Howard et al. utilized a MobileNet for mobile and embedded vision applications.

Figure 3: *Descriptions of models and sources*

3.2. Measuring Performance

To evaluate the performance of the models, we compared their performance on a test set using metrics such as accuracy, recall, precision, and F-score (F1), defined as:

$$Accuracy = \frac{True\ predictions}{Size\ of\ set} \qquad Precision\ of\ a\ category = \frac{True\ predictions\ of\ the\ category}{Size\ of\ predicted\ set\ of\ that\ category}$$

$$Recall\ of\ a\ category = \frac{True\ predictions\ of\ the\ category}{True\ set\ size\ of\ that\ category} \qquad F1 = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}}$$

where *True Predictions* is defined as all correct predictions in the 4 categories.

4. Data Processing and Augmentation

4.1. Data Augmentation

We added flipping transformation to the images to increase the robustness of the model as images might sometimes be oriented differently, using RandomHorizontalFlip and RandomVerticalFlip. Furthermore, since our dataset included images with different dimensions, the images were standardized before the models were trained. We thus resized the images using the Resize function of torch transformations (where we used the smallest image dimensions, 64 x 64), by using `transforms.Resize(size=(64, 64))`. Based on existing research that involves detecting carrot diseases using CNNs [19], we also found that we can use the Gaussian Blur transformation to remove Gaussian Noise to minimize image noise from different sources without hampering other features. However, this transformation was already applied when we resized the images, meaning we did not need to reapply it.

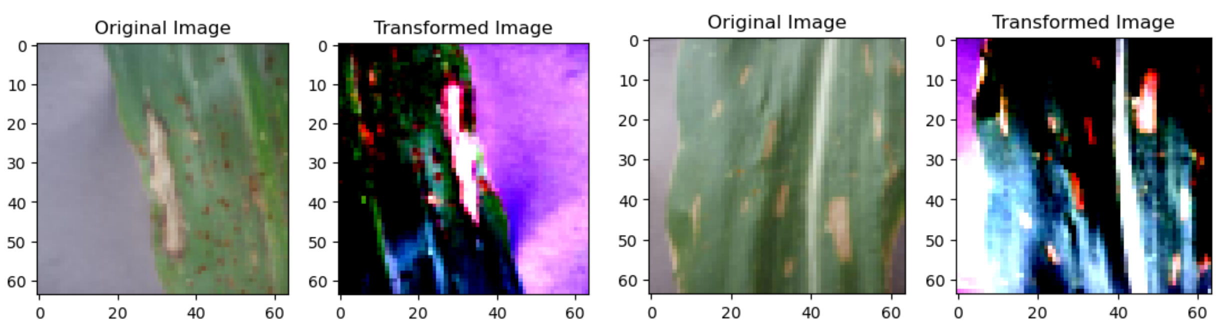


Figure 4: *Images of corn leaf blight before and after data augmentation is applied*

4.2. Problems With the Dataset

After initially inspecting the different pictures in our dataset, a certain problem became very visible. As shown in the figure below, the type of leaf in a picture can easily be identified based on the color of the background: leaves with blight or gray leaf spots have a light purple background, leaves with common rust have a black background and healthy leaves are mostly zoomed in and do not have a background. This relationship indicates that the color of the background could end up being used by the CNN model, making it highly dependent on the wrong features of the image and ultimately useless–a hypothesis we examined and confirmed later on in the research process. We considered several approaches to fix this problem: using grayscale versions of the images (which we rapidly discarded as the diseases can be recognized by the color of the leaf and we would lose a lot of information), zooming in enough to eliminate the background, or standardizing the background color across all diseases.

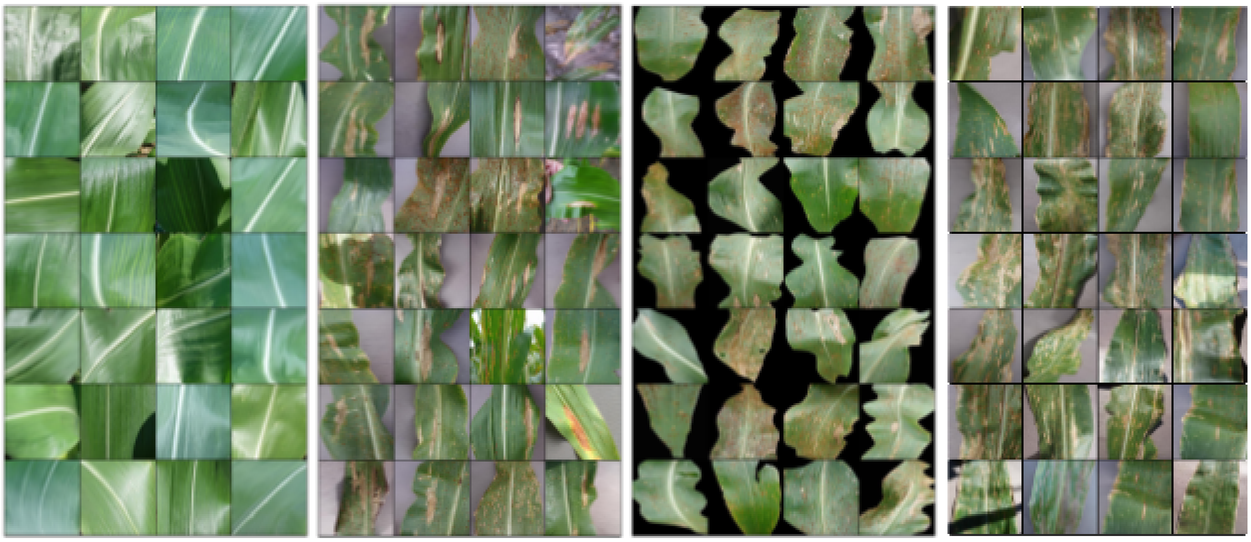


Figure 5*: 28 randomly selected "healthy", "blight", "common rust" and "gray leaf spot" leave pictures*

We chose the last option as the visible indicators of diseases on the leaves vary in their location from leaf to leaf, necessitating manual zooming on each leaf. Given the impracticality of individually processing images in a large dataset, we decided to remove the backgrounds from all images and replace them with a black background.

4.3. Background Removal

We tried several methods to remove the backgrounds of the leaves. The first involved using the existing Segment Anything AI model from Meta AI [20] to automate background removal with no help. Although the model successfully removed the backgrounds of most images, around 1/5 of the images still had most of their background, as shown below.

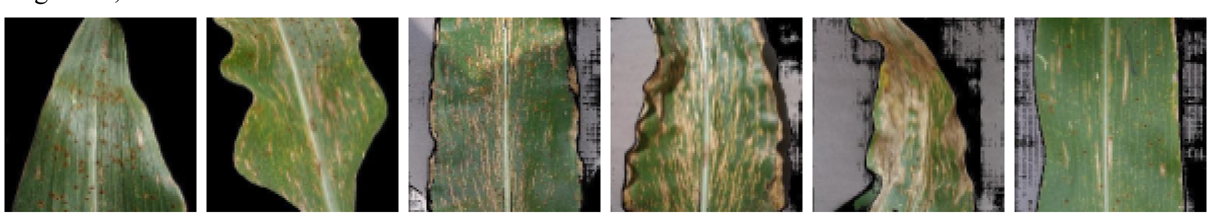


Figure 6*: 2 successfully removed backgrounds and 4 unsuccessfully removed backgrounds using Segment Anything, all pictures out of the "Gray Leaf Spot" category*

This was likely due to the non-uniform backgrounds and the leaves covering a large portion of the picture, which may have caused the model to struggle in determining where to focus within the image.

The second attempt involved edge detection, for which we tried using the OpenCV library to implement the Sobel Edge detection algorithm. However, the fact that the leaf wasn't entirely in the picture caused trouble, as the model couldn't detect the edge of the leaves connected to the sides of the picture. We tried adding light purple padding to the image, but because the model's background wasn't uniform, we couldn't find a sensitivity where it wouldn't differentiate between the padding, background, and sides of the leaf. In the image below, the algorithm correctly finds the leaf's edges but fails to convert it to the wanted contours since the left, top, and bottom sides of the leaf aren't completely visible and thus it does not detect an edge there. Although for this specific image, it would be possible to make an algorithm connect two edges at the top and the bottom, this method would not work more generally because leaves could have no edge on the left side (i.e. be completely out of frame on the left).

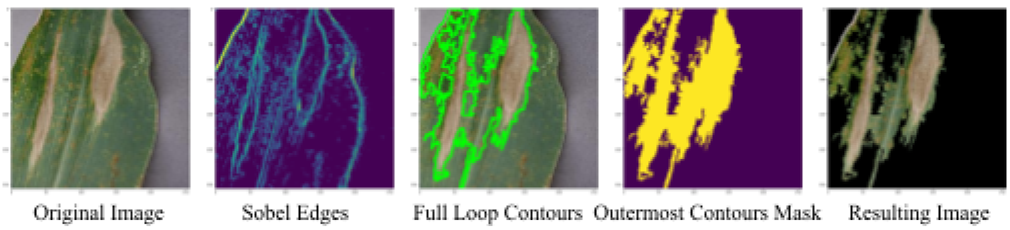


***Figure** 7: Example of the background removal process of a "Gray Leaf Spot" image using edge detection*

We thus decided to use color-based segmentation to mask green, brown, and yellow pixels in the image and remove any other pixels. This was also done with OpenCV by first converting the pictures from RGB to HSV before creating a temporary mask. The mask was then converted to contours with the Segment Anything method's padding technique: we took the mask, added 20 pixels of padding around the image, then took the outermost contours, removed the 20 pixels of padding, and converted it into a mask.

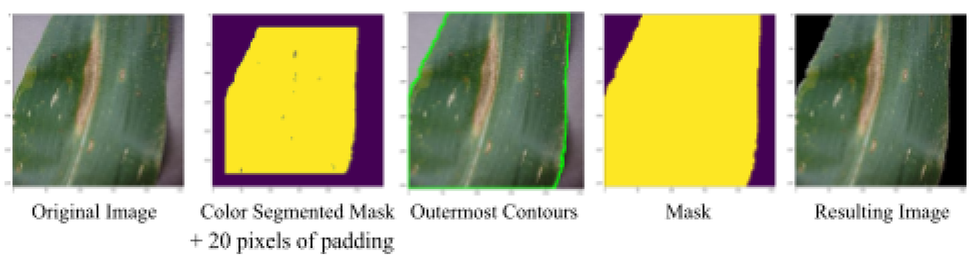


***Figure 8**: Example of the background removal process of a "Gray Leaf Spot" image using color-based segmentation*

5 Models

5.1. Preliminary Model Performances

In order to see how much the models would overfit the backgrounds, we first trained the models on the original images. This section will be looking at those models exclusively.
The trained models are described in Figure 3. They were trained using 30 Epochs, and L2 Regularization with a training set of 70% of the original dataset and a validation set of 20%. The models were then analyzed with the test set, which contains the remaining 10% of the dataset.

Original model performances on the test set were strong, as shown by the first column below. However, when we added pictures with backgrounds that did not align with the patterns illustrated in Figure 5 (those images will be referred to as images with "different backgrounds"), accuracies dropped significantly, indicating that our models heavily overfitted to the backgrounds of the images:

Model	Accuracy on test set	Accuracy on different background test set	Total Accuracy (weighted average of both)
ResNet 50	0.841	0.262	0.649
ResNet 101	0.807	0.357	0.635
ResNet 152	0.820	0.371	0.671
VGG16	0.818	0.347	0.642
EfficientNet	0.890	0.490	0.738
MobileNet	0.891	0.506	0.744
DenseNet	0.906	0.433	0.725
LeNet	0.862	0.283	0.645
ShuffleNet	0.908	0.440	0.728

Figure 9*: Accuracies of models trained on the original dataset*

Although certain models performed relatively well on the test set, their poor performance on the test set with different backgrounds prevents them from becoming viable implementations of the technology—a model that can only make good predictions when the background gives away the right classification is not useful.

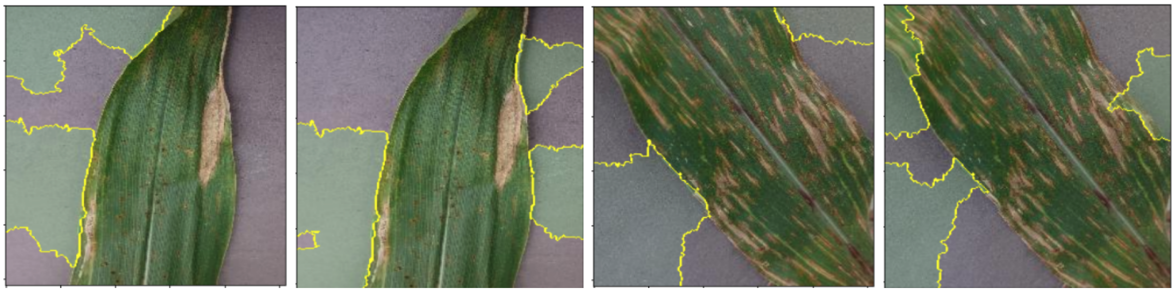
The models' difficulty with gray leaf spot images, indicated by the recall for this category, also indicates overfitting as these images were commonly misclassified as blight, which has similarly colored backgrounds in the dataset (see Figure 5).

ResNet 50	ResNet 101	ResNet 152	VGG 16	Efficient Net	Mobile Net	Dense Net	LeNet	Shuffle Net
0.482	0.000	0.156	0.156	0.298	0.418	0.440	0.206	0.440

Figure 10: *Recall for the Gray Leaf Spot category of the original test set*

5.2. Understanding Models

Deconvolution tools such as LIME (Local Interpretable Model-Agnostic Explanations) [21] [22] make the problem even more apparent. By repeatedly altering some specific parts of the image, the tool can find the areas of the picture that influence the model's decision the most. These areas are then given a yellow boundary and highlighted with green for emphasis. The results speak for themselves. As shown in the following pictures, the models use the background extensively in their decision-making processes, leading them to wrongly classify the image on the right as blight instead of gray leaf spot.



ShuffleNet **MobileNet** **LeNet** **EfficientNet**

Figure 11: *LIME applied to two images for two different models each.*

The models seem to struggle even more with the pictures without the standard backgrounds (purple for blight and gray leaf spot, black for common rust, and none for healthy leaves). EfficientNet predicted the leaf on the right to have blight, although it has visible rust. The image on the right was classified as containing rust by the model, although it is completely healthy.

Predicted: Blight | Actual: Common Rust **Predicted: Common Rust | Actual: Healthy**

Figure 12: *LIME applied to two images with different backgrounds*

5.3. Re-Trained Model Performances

Following image segmentation, we retrained our models using the same parameters. Note that as all backgrounds were now standardized, there was no point in testing the models on the images with different backgrounds. We thus examined their overall accuracy exclusively.

ResNet 50	ResNet 101	ResNet 152	VGG 16	Efficient Net	Mobile Net	Dense Net	LeNet	Shuffle Net
0.787	0.838	0.782	0.758	0.804	0.877	0.891	0.827	0.874
0.138▲	0.203▲	0.111▲	0.116▲	0.066▲	0.133▲	0.166▲	0.182▲	0.146▲

Figure 13: *Accuracy of the models after training on a segmented dataset*

As shown in the figure above, all models improved after training on the segmented dataset, with ResNet101 getting the biggest improvement of 20.3%, and DenseNet reaching the highest accuracy of 89.1%.

5.4. Understanding New Models

MobileNetV1 LeNet ShuffleNet EfficientNet Lenet

Figure 14: Lime applied to different models and segmented images

Again, we used Lime to better understand the new models' performance. As seen in the images above, the new models learned to focus exclusively on features of the leaf to correctly classify the images, as the background was standardized and provided no information that would be helpful. Removing the background from the model's decision-making process also allowed it to do a better job distinguishing between blight and gray leaf spot, as shown in the last picture above.

Additionally, we can directly compare models before and after training on segmented backgrounds. When put side by side in the table below, the improvement is obvious: all models trained on segmented backgrounds use parts of the leaves exclusively.

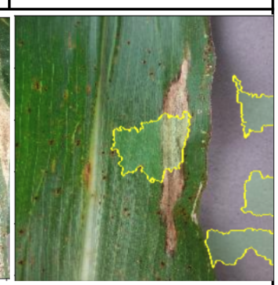
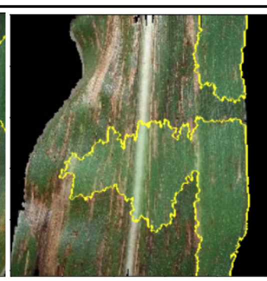
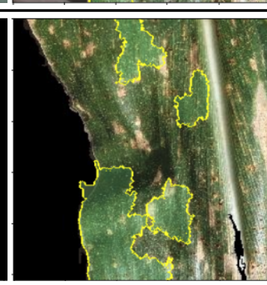
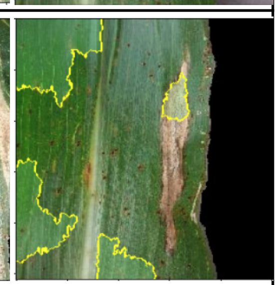
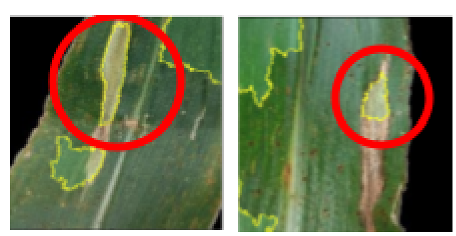
Model	LeNet	EfficientNet	ShuffleNet	MobileNet
Before segmentation				
After segmentation				

Figure 15: Understand model performance before versus after segmentation with Lime

Furthermore, upon closer inspection, the regions highlighted by LIME include the parts of the leaf where the disease is the most visible, indicating that the model is using the correct features of the images to make its decisions. This is another improvement over the original models.



6. Conclusion and Future Work

In this study, we sought to create effective models for disease classification in corn leaves while also addressing the challenges posed by the Maize Leaf Disease dataset. We experimented with various neural network architectures, including ResNet, VGG16, EfficientNet, MobileNet, DenseNet, LeNet, and ShuffleNet. However, this dataset presented a significant issue: images were categorized by disease with distinct backgrounds unique to each category. Consequently, our initial models overly relied on these backgrounds, resulting in poor generalization when the backgrounds were not standard.

To address this, we standardized the backgrounds of the images, ensuring the models focused on the leaves themselves. We tried several methods and decided on using color-based segmentation, which returned the most consistent results. Our newly trained models demonstrated a notable improvement in model performance once the backgrounds were removed, as evidenced by higher accuracies and recall across the board. This improvement was especially visible when the deconvolution tool LIME was used, as we were able to confirm that the models trained on standardized backgrounds relied on relevant leaf features rather than extraneous background information.

In conclusion, this study highlights the critical role of a large and diverse dataset and proper preprocessing in training effective deep-learning models. By standardizing the backgrounds, we ensured the models could generalize better and make accurate predictions based on the actual diseased areas of the corn leaves. This approach can be extended to other agricultural datasets and diseases, paving the way for more reliable and efficient disease detection systems in the future.

As a next step, we could review the architecture of the models we used, experimenting with different optimization and L2 regularization. Furthermore, we could also consider using transfer learning of existing agricultural models that worked for other types of leaves, as this would deal with the fact that the dataset we used is too small. It could also help with overfitting to the backgrounds. If possible, however, the best solution would probably involve expanding the dataset and incorporating more diverse backgrounds to enhance model robustness in real-world scenarios.

7. References

[1] Mueller, Daren, et al. "Corn Disease Loss Estimates from the United States and Ontario, Canada - 2022." Crop Protection Network, 6 Feb. 2023.

[2] Saavoss, Monica, et al. "Trends in Production Practices and Costs of the U.S. Corn Sector." Economic Research Report Number 294, USDA Economic Research Service, U.S. Department of Agriculture, July 2021

[3] "Common Rust." *Cornell College of Agriculture and Life Sciences*, CALS.

[4] Jackson-Ziems, Tamara A. "Rust Diseases of Corn in Nebraska." Nebraska Extension Publications, University of Nebraska-Lincoln, Jan. 2014.

[5] Malvick, Dean. "Gray Leaf Spot on Corn." UMN Extension, University of Minnesota, 2018.

[6] "Gray Leaf Spot." *Cornell College of Agriculture and Life Sciences*, CALS.

[7] Wise, Kiersten. "Gray Leaf Spot." Purdue Extension, Purdue University.

[8] Malvick, Dean. "Gray Leaf Spot on Corn." UMN Extension, University of Minnesota, 2018.

[9] Wise, Kiersten. "Northern Corn Leaf Blight." Purdue Extension, Purdue University

[10] Ghose, Smaranjit. "Corn or Maize Leaf Disease Dataset." Artificial Intelligence Based Classification of Diseases in Maize/Corn Plants, 11 Nov. 2020.

[11] Paszke, Adam, et al. 'PyTorch: An Imperative Style, High-Performance Deep Learning Library'. Advances in Neural Information Processing Systems 32, Curran Associates, Inc., 2019, pp. 8024–8035.

[12] Godi, Brahmaji et al. "ResNet Model to Forecast Plant Leaf Disease." *Computer Science & Engineering*, Raghu Institute of Technology, Visakhapatnam, India, 2024. *IEEE Xplore*.

[13] Bhatt, Prakruti V. et al. "Identification of Diseases in Corn Leaves using Convolutional Neural Networks and Boosting." International Conference on Pattern Recognition Applications and Methods (2019).

[14] Sun, Xuewei, et al. "Research on plant disease identification based on CNN." Cognitive Robotics, vol. 2, 2022, pp. 155–163.

[15] Wei, Guangfen et al. "Development of a LeNet-5 Gas Identification CNN Structure for Electronic Noses." Sensors (Basel, Switzerland) vol. 19,1 217. 8 Jan. 2019, doi:10.3390/s19010217

[16] Zhang, Xiangyue, et al. "ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices." ArXiv, 7 Dec. 2017, arxiv.org/pdf/1707.01083

[17] Huang, Gao, et al. "Densely Connected Convolutional Networks." arXiv preprint arXiv:1608.06993 (2018).

[18] Howard, Andrew G, et al. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." ArXiv, 17 Apr 2017, arxiv.org/pdf/1704.04861.

[19] Ray, Shree. Dolax, et al. 'Carrot Cure: A CNN Based Application to Detect Carrot Disease'. 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), IEEE, 2022, pp. 01–07. DOI.org (Crossref).

[20] Kirillov, Alexander, et al. "Segment Anything." ArXiv, Meta AI Research, FAIR, 5 Apr. 2023, arxiv.org/pdf/2304.02643.pdf.

[21] Ribeiro, Marco, et al. "Why Should I Trust You?" Explaining the Predictions of Any Classifier. ArXiv, 16 Feb 2016, arxiv.org/pdf/1602.04938v1.

[22] Ribeiro, Marco, et al. "Local Interpretable Model-Agnostic Explanations (LIME): An Introduction." O'Reilly Media, 12 Aug. 2016, www.oreilly.com/content/introduction-to-local-interpretable-model-agnostic-explanations-lime/.

8. Appendix

8.1 Confusion Matrices for Models Trained on Non-Segmented Backgrounds, Test Set

ResNet50

Actual

Blight	96	0	19	0
Common_Rust	10	117	4	0
Gray_Leaf_Spot	25	1	33	0
Healthy	8	0	0	109
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

ResNet101

Actual

Blight	113	0	1	2
Common_Rust	11	116	2	4
Gray_Leaf_Spot	57	0	0	2
Healthy	3	0	0	114
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

ResNet152

Actual

Blight	89	11	11	4
Common_Rust	3	126	1	1
Gray_Leaf_Spot	26	17	15	1
Healthy	0	1	0	116
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

VGG16

Actual

Blight	110	1	3	1
Common_Rust	9	115	6	1
Gray_Leaf_Spot	45	5	8	1
Healthy	4	1	0	112
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

EfficientNet

Actual	Blight	109	5	1	0
	Common_Rust	4	123	1	1
	Gray_Leaf_Spot	27	7	25	0
	Healthy	0	0	0	117
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

MobileNet

Actual	Blight	100	3	12	0
	Common_Rust	6	124	1	0
	Gray_Leaf_Spot	13	10	36	0
	Healthy	0	1	0	116
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

DenseNet

Actual	Blight	110	1	4	1
	Common_Rust	5	126	0	2
	Gray_Leaf_Spot	22	3	33	1
	Healthy	1	0	0	116
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

LeNet

Actual	Blight	112	0	2	1
	Common_Rust	10	119	2	0
	Gray_Leaf_Spot	36	5	18	0
	Healthy	1	1	0	113
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

ShuffleNet

Actual	Blight	105	0	0	1
	Common_Rust	7	122	1	1
	Gray_Leaf_Spot	24	2	32	1
	Healthy	1	0	0	116
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

8.2 Performance Metrics for Models Trained on Non-Segmented Backgrounds, Test Set

ResNet50

Accuracy:	0.841				
Precision for Blight:	0.691	**Recall for Blight:**	0.835	**F1 Blight:**	0.756
Precision for Common Rust:	0.992	**Recall for Common Rust:**	0.893	**F1 Common Rust:**	0.940
Precision for Gray Leaf Spot:	0.589	**Recall for Gray Leaf Spot:**	0.559	**F1 Gray Leaf Spot:**	0.574
Precision for Healthy:	1.000	**Recall for Healthy:**	0.932	**F1 Healthy:**	0.965

ResNet101

Accuracy:	0.807				
Precision for Blight:	0.614	**Recall for Blight:**	0.974	**F1 Blight:**	0.753
Precision for Common Rust:	1.000	**Recall for Common Rust:**	0.872	**F1 Common Rust:**	0.932
Precision for Gray Leaf Spot:	0.000	**Recall for Gray Leaf Spot:**	0.000	**F1 Gray Leaf Spot:**	0.000
Precision for Healthy:	0.934	**Recall for Healthy:**	0.974	**F1 Healthy:**	0.954

ResNet152

Accuracy:	0.820				
Precision for Blight:	0.754	**Recall for Blight:**	0.774	**F1 Blight:**	0.764
Precision for Common Rust:	0.813	**Recall for Common Rust:**	0.962	**F1 Common Rust:**	0.881
Precision for Gray Leaf Spot:	0.556	**Recall for Gray Leaf Spot:**	0.254	**F1 Gray Leaf Spot:**	0.349
Precision for Healthy:	0.951	**Recall for Healthy:**	0.991	**F1 Healthy:**	0.971

VGG16

Accuracy:	0.818				
Precision for Blight:	0.655	**Recall for Blight:**	0.957	**F1 Blight:**	0.777
Precision for Common Rust:	0.943	**Recall for Common Rust:**	0.878	**F1 Common Rust:**	0.909
Precision for Gray Leaf Spot:	0.471	**Recall for Gray Leaf Spot:**	0.136	**F1 Gray Leaf Spot:**	0.211
Precision for Healthy:	0.974	**Recall for Healthy:**	0.957	**F1 Healthy:**	0.966

EfficientNet

Accuracy:	0.890				
Precision for Blight:	0.779	**Recall for Blight:**	0.948	**F1 Blight:**	0.855
Precision for Common Rust:	0.911	**Recall for Common Rust:**	0.953	**F1 Common Rust:**	0.932
Precision for Gray Leaf Spot:	0.926	**Recall for Gray Leaf Spot:**	0.424	**F1 Gray Leaf Spot:**	0.581
Precision for Healthy:	0.992	**Recall for Healthy:**	1.000	**F1 Healthy:**	0.996

MobileNet

Accuracy:	0.891				
Precision for Blight:	0.840	**Recall for Blight:**	0.870	**F1 Blight:**	0.855
Precision for Common Rust:	0.899	**Recall for Common Rust:**	0.947	**F1 Common Rust:**	0.922
Precision for Gray Leaf Spot:	0.735	**Recall for Gray Leaf Spot:**	0.610	**F1 Gray Leaf Spot:**	0.667
Precision for Healthy:	1.000	**Recall for Healthy:**	0.991	**F1 Healthy:**	0.996

DenseNet

Accuracy:	0.906				
Precision for Blight:	0.797	**Recall for Blight:**	0.948	**F1 Blight:**	0.866
Precision for Common Rust:	0.969	**Recall for Common Rust:**	0.947	**F1 Common Rust:**	0.958
Precision for Gray Leaf Spot:	0.892	**Recall for Gray Leaf Spot:**	0.559	**F1 Gray Leaf Spot:**	0.688
Precision for Healthy:	0.967	**Recall for Healthy:**	0.991	**F1 Healthy:**	0.979

LeNet

Accuracy:	0.862				
Precision for Blight:	0.704	**Recall for Blight:**	0.974	**F1 Blight:**	0.818
Precision for Common Rust:	0.952	**Recall for Common Rust:**	0.908	**F1 Common Rust:**	0.930
Precision for Gray Leaf Spot:	0.818	**Recall for Gray Leaf Spot:**	0.305	**F1 Gray Leaf Spot:**	0.444
Precision for Healthy:	0.991	**Recall for Healthy:**	0.983	**F1 Healthy:**	0.987

ShuffleNet

Accuracy:	0.908				
Precision for Blight:	0.766	**Recall for Blight:**	0.991	**F1 Blight:**	0.864
Precision for Common Rust:	0.984	**Recall for Common Rust:**	0.931	**F1 Common Rust:**	0.957
Precision for Gray Leaf Spot:	0.970	**Recall for Gray Leaf Spot:**	0.542	**F1 Gray Leaf Spot:**	0.696
Precision for Healthy:	0.975	**Recall for Healthy:**	0.991	**F1 Healthy:**	0.983

8.3 Confusion Matrices for Models Trained on Non-Segmented Backgrounds, Different Background Set

ResNet50

Actual (rows) / Predicted (columns)

Blight	1	0	0	0
Common_Rust	51	3	15	1
Gray_Leaf_Spot	44	1	35	2
Healthy	37	1	3	16
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

ResNet101

Blight	48	1	2	3
Common_Rust	51	3	3	13
Gray_Leaf_Spot	71	1	0	10
Healthy	14	0	0	43
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

ResNet152

Blight	0	1	0	0
Common_Rust	14	33	5	18
Gray_Leaf_Spot	19	48	7	8
Healthy	4	15	0	38
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

VGG16

Blight	27	5	17	1
Common_Rust	41	15	6	0
Gray_Leaf_Spot	38	22	14	8
Healthy	16	8	2	31
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

EfficientNet

Actual					
	Blight	43	7	0	0
	Common_Rust	22	36	2	10
Actual	Gray_Leaf_Spot	43	19	17	3
	Healthy	17	9	0	31
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

MobileNet

	Blight	37	7	6	0
	Common_Rust	22	45	2	1
Actual	Gray_Leaf_Spot	26	31	23	2
	Healthy	20	11	0	26
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

DenseNet

	Blight	43	7	4	0
	Common_Rust	21	35	7	7
Actual	Gray_Leaf_Spot	38	12	29	3
	Healthy	17	33	0	7
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

LeNet

	Blight	42	3	5	0
	Common_Rust	45	18	6	1
Actual	Gray_Leaf_Spot	57	12	11	2
	Healthy	43	4	2	0
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

ShuffleNet

Actual					
	Blight	42	4	4	0
	Common_Rust	24	33	8	5
	Gray_Leaf_Spot	41	8	30	3
	Healthy	15	33	0	9
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

8.4 Performance Metrics for Models Trained on Non-Segmented Backgrounds, Different Background Set

ResNet50

Accuracy:	0.262				
Precision for Blight:	0.008	**Recall for Blight:**	1.000	**F1 Blight:**	0.015
Precision for Common Rust:	0.600	**Recall for Common Rust:**	0.043	**F1 Common Rust:**	0.080
Precision for Gray Leaf Spot:	0.660	**Recall for Gray Leaf Spot:**	0.427	**F1 Gray Leaf Spot:**	0.519
Precision for Healthy:	0.842	**Recall for Healthy:**	0.281	**F1 Healthy:**	0.421

ResNet101

Accuracy:	0.357				
Precision for Blight:	0.261	**Recall for Blight:**	0.889	**F1 Blight:**	0.403
Precision for Common Rust:	0.600	**Recall for Common Rust:**	0.043	**F1 Common Rust:**	0.080
Precision for Gray Leaf Spot:	0.000	**Recall for Gray Leaf Spot:**	0.000	**F1 Gray Leaf Spot:**	0.000
Precision for Healthy:	0.623	**Recall for Healthy:**	0.754	**F1 Healthy:**	0.683

ResNet152

Accuracy:	0.371				
Precision for Blight:	0.000	**Recall for Blight:**	0.000	**F1 Blight:**	0.000
Precision for Common Rust:	0.340	**Recall for Common Rust:**	0.471	**F1 Common Rust:**	0.395
Precision for Gray Leaf Spot:	0.583	**Recall for Gray Leaf Spot:**	0.085	**F1 Gray Leaf Spot:**	0.149
Precision for Healthy:	0.594	**Recall for Healthy:**	0.667	**F1 Healthy:**	0.628

VGG16

Accuracy:	0.347				
Precision for Blight:	0.221	**Recall for Blight:**	0.540	**F1 Blight:**	0.314
Precision for Common Rust:	0.300	**Recall for Common Rust:**	0.242	**F1 Common Rust:**	0.268
Precision for Gray Leaf Spot:	0.359	**Recall for Gray Leaf Spot:**	0.171	**F1 Gray Leaf Spot:**	0.231
Precision for Healthy:	0.775	**Recall for Healthy:**	0.544	**F1 Healthy:**	0.639

EfficientNet

Accuracy:	0.490				
Precision for Blight:	0.344	**Recall for Blight:**	0.860	**F1 Blight:**	0.491
Precision for Common Rust:	0.507	**Recall for Common Rust:**	0.514	**F1 Common Rust:**	0.511
Precision for Gray Leaf Spot:	0.895	**Recall for Gray Leaf Spot:**	0.207	**F1 Gray Leaf Spot:**	0.337
Precision for Healthy:	0.705	**Recall for Healthy:**	0.544	**F1 Healthy:**	0.614

MobileNet

Accuracy:	0.506				
Precision for Blight:	0.352	**Recall for Blight:**	0.740	**F1 Blight:**	0.477
Precision for Common Rust:	0.479	**Recall for Common Rust:**	0.643	**F1 Common Rust:**	0.549
Precision for Gray Leaf Spot:	0.742	**Recall for Gray Leaf Spot:**	0.280	**F1 Gray Leaf Spot:**	0.407
Precision for Healthy:	0.897	**Recall for Healthy:**	0.456	**F1 Healthy:**	0.605

DenseNet

Accuracy:	0.433				
Precision for Blight:	0.361	**Recall for Blight:**	0.796	**F1 Blight:**	0.497
Precision for Common Rust:	0.402	**Recall for Common Rust:**	0.500	**F1 Common Rust:**	0.446
Precision for Gray Leaf Spot:	0.725	**Recall for Gray Leaf Spot:**	0.354	**F1 Gray Leaf Spot:**	0.475
Precision for Healthy:	0.412	**Recall for Healthy:**	0.123	**F1 Healthy:**	0.189

LeNet

Accuracy:	0.283				
Precision for Blight:	0.225	**Recall for Blight:**	0.840	**F1 Blight:**	0.354
Precision for Common Rust:	0.486	**Recall for Common Rust:**	0.257	**F1 Common Rust:**	0.336
Precision for Gray Leaf Spot:	0.458	**Recall for Gray Leaf Spot:**	0.134	**F1 Gray Leaf Spot:**	0.208
Precision for Healthy:	0.000	**Recall for Healthy:**	0.000	**F1 Healthy:**	0.000

ShuffleNet

Accuracy:	0.440				
Precision for Blight:	0.344	**Recall for Blight:**	0.840	**F1 Blight:**	0.488
Precision for Common Rust:	0.423	**Recall for Common Rust:**	0.471	**F1 Common Rust:**	0.446
Precision for Gray Leaf Spot:	0.714	**Recall for Gray Leaf Spot:**	0.366	**F1 Gray Leaf Spot:**	0.484
Precision for Healthy:	0.529	**Recall for Healthy:**	0.158	**F1 Healthy:**	0.243

8.5 Confusion Matrices for Models Trained on Segmented Backgrounds

ResNet50

Actual

Blight	99	10	6	0
Common_Rust	21	103	7	0
Gray_Leaf_Spot	37	6	16	0
Healthy	2	1	0	114
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

ResNet101

Actual

Blight	96	11	5	1
Common_Rust	7	124	0	0
Gray_Leaf_Spot	35	8	16	0
Healthy	1	0	0	116
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

ResNet152

Actual

Blight	82	13	17	3
Common_Rust	13	111	6	1
Gray_Leaf_Spot	27	9	23	0
Healthy	1	2	0	114
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

VGG16

Actual

Blight	95	16	0	4
Common_Rust	17	108	0	6
Gray_Leaf_Spot	43	14	0	2
Healthy	0	0	0	117
	Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

EfficientNet

Actual	Blight	89	1	20	5
	Common_Rust	2	97	7	25
	Gray_Leaf_Spot	6	2	41	16
	Healthy	0	0	0	117
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

MobileNet

Actual	Blight	106	4	4	1
	Common_Rust	9	119	3	0
	Gray_Leaf_Spot	19	11	29	0
	Healthy	0	1	0	116
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

DenseNet

Actual	Blight	100	4	10	1
	Common_Rust	7	118	5	1
	Gray_Leaf_Spot	12	4	43	0
	Healthy	1	1	0	115
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

LeNet

Actual	Blight	113	1	0	1
	Common_Rust	11	120	0	0
	Gray_Leaf_Spot	50	9	0	0
	Healthy	1	0	0	116
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

ShuffleNet

Actual	Blight	98	6	10	1
	Common_Rust	8	120	3	0
	Gray_Leaf_Spot	22	3	34	0
	Healthy	0	0	0	117
		Blight	Common_Rust	Gray_Leaf_Spot	Healthy

Predicted

8.6 Performance Metrics for Models Trained on Segmented Backgrounds

ResNet50

Accuracy:	0.787				
Precision for Blight:	0.623	**Recall for Blight:**	0.861	**F1 Blight:**	0.723
Precision for Common Rust:	0.858	**Recall for Common Rust:**	0.786	**F1 Common Rust:**	0.821
Precision for Gray Leaf Spot:	0.552	**Recall for Gray Leaf Spot:**	0.271	**F1 Gray Leaf Spot:**	0.364
Precision for Healthy:	1.000	**Recall for Healthy:**	0.974	**F1 Healthy:**	0.987

ResNet101

Accuracy:	0.838				
Precision for Blight:	0.691	**Recall for Blight:**	0.850	**F1 Blight:**	0.762
Precision for Common Rust:	0.867	**Recall for Common Rust:**	0.947	**F1 Common Rust:**	0.905
Precision for Gray Leaf Spot:	0.762	**Recall for Gray Leaf Spot:**	0.271	**F1 Gray Leaf Spot:**	0.400
Precision for Healthy:	0.991	**Recall for Healthy:**	0.991	**F1 Healthy:**	0.991

ResNet152

Accuracy:	0.782				
Precision for Blight:	0.667	**Recall for Blight:**	0.713	**F1 Blight:**	0.689
Precision for Common Rust:	0.822	**Recall for Common Rust:**	0.847	**F1 Common Rust:**	0.835
Precision for Gray Leaf Spot:	0.500	**Recall for Gray Leaf Spot:**	0.390	**F1 Gray Leaf Spot:**	0.438
Precision for Healthy:	0.966	**Recall for Healthy:**	0.974	**F1 Healthy:**	0.970

VGG16

Accuracy:	0.758				
Precision for Blight:	0.613	**Recall for Blight:**	0.826	**F1 Blight:**	0.704
Precision for Common Rust:	0.783	**Recall for Common Rust:**	0.824	**F1 Common Rust:**	0.803
Precision for Gray Leaf Spot:	0.000	**Recall for Gray Leaf Spot:**	0.000	**F1 Gray Leaf Spot:**	0.000
Precision for Healthy:	0.907	**Recall for Healthy:**	1.000	**F1 Healthy:**	0.951

EfficientNet

Accuracy:	0.804				
Precision for Blight:	0.918	**Recall for Blight:**	0.774	**F1 Blight:**	0.840
Precision for Common Rust:	0.970	**Recall for Common Rust:**	0.740	**F1 Common Rust:**	0.840
Precision for Gray Leaf Spot:	0.603	**Recall for Gray Leaf Spot:**	0.631	**F1 Gray Leaf Spot:**	0.617
Precision for Healthy:	0.718	**Recall for Healthy:**	1.000	**F1 Healthy:**	0.836

MobileNet

Accuracy:	0.877				
Precision for Blight:	0.791	**Recall for Blight:**	0.922	**F1 Blight:**	0.851
Precision for Common Rust:	0.881	**Recall for Common Rust:**	0.908	**F1 Common Rust:**	0.895
Precision for Gray Leaf Spot:	0.806	**Recall for Gray Leaf Spot:**	0.492	**F1 Gray Leaf Spot:**	0.611
Precision for Healthy:	0.991	**Recall for Healthy:**	0.991	**F1 Healthy:**	0.991

DenseNet

Accuracy:	0.891				
Precision for Blight:	0.833	**Recall for Blight:**	0.870	**F1 Blight:**	0.851
Precision for Common Rust:	0.929	**Recall for Common Rust:**	0.901	**F1 Common Rust:**	0.915
Precision for Gray Leaf Spot:	0.741	**Recall for Gray Leaf Spot:**	0.729	**F1 Gray Leaf Spot:**	0.735
Precision for Healthy:	0.983	**Recall for Healthy:**	0.983	**F1 Healthy:**	0.983

LeNet

Accuracy:	0.827				
Precision for Blight:	0.646	**Recall for Blight:**	0.983	**F1 Blight:**	0.779
Precision for Common Rust:	0.923	**Recall for Common Rust:**	0.916	**F1 Common Rust:**	0.920
Precision for Gray Leaf Spot:	0.000	**Recall for Gray Leaf Spot:**	0.000	**F1 Gray Leaf Spot:**	0.000
Precision for Healthy:	0.991	**Recall for Healthy:**	0.991	**F1 Healthy:**	0.991

ShuffleNet

Accuracy:	0.874				
Precision for Blight:	0.766	**Recall for Blight:**	0.852	**F1 Blight:**	0.807
Precision for Common Rust:	0.930	**Recall for Common Rust:**	0.916	**F1 Common Rust:**	0.923
Precision for Gray Leaf Spot:	0.723	**Recall for Gray Leaf Spot:**	0.576	**F1 Gray Leaf Spot:**	0.642
Precision for Healthy:	0.992	**Recall for Healthy:**	1.000	**F1 Healthy:**	0.996