PHOTOGRAPHIC DIAGNOSIS OF PECAN LEAF DISEASE

An Undergraduate Research Scholars Thesis

by

HARRISON FROESCHKE

Submitted to the Undergraduate Research Scholars program at
Texas A&M University
in partial fulfillment of the requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by Research Advisor: Dr. Bruce Gooch

May 2018

Major: Computer Science

TABLE OF CONTENTS

ABSTRACT

Photographic Diagnosis of Pecan Leaf Disease

Harrison Froeschke
Department of Computer Science
Texas A&M University

Research Advisor: Dr. Bruce Gooch
Department of Computer Science
Texas A&M University

Ever since Texas Governor James Stephen Hogg in 1906 requested that a pecan tree serves as his headstone and that the nuts of the tree be distributed to make Texas a "Land of Trees," pecans have been a big part of many Texan's identities. So much so did Texans love pecans that the pecan tree became the official tree of Texas in 1919 in the context of the 36th Texas Legislature. [1] *Fusicladium effusum*, commonly called pecan scab, is the most economically impactful disease of trees in the south-eastern United States. In 2013, there was an estimated 20-25 million pound loss of pecans due to weather conditions that allowed pecan scab to thrive. [2]. The fungus is a problem for pecan growers as it reduces both yield and profit. Our work will aid researchers and farmers in identifying and monitoring pecan scab infections in their orchards by analysing a photo of a pecan tree leaf taken by the user with a mobile phone. Having this tool will allow farmers to react faster to pecan scab, have a good idea of how much fungicide is appropriate to deal with the level of infection, and compare disease severity over the time the user has sampled the pecan leaves.

DEDICATION

This paper is dedicated to the researchers of the Noble Research Institute who work hard to improve agriculture in the U.S. and constantly strive for greatness, the state of agriculture across the U.S., and the South would be leagues behind without their help.

Also, this paper is dedicated to Dr. Bruce Gooch, who believed in me enough to give me a spot researching with him in his lab and constantly supports me in whatever I do. Thank you for guiding me through my years at Texas A&M and always entertaining me with your great stories.

ACKNOWLEDGEMENTS

I would like to thank Dr. Charles Rohla and Dr. Stephen Webb for being my contacts at the Noble Research Institute, giving us the prompt and permission for this project, valuable data that is needed for this project's completion, and being experts in their field to help on the biology side of this project.

A large amount of thanks goes to Dr. Bruce Gooch, who provided me guidance with the project from the beginning and especially when I had to complete the project on my own. Without Dr. Gooch, this paper and research would not be possible.

NOMENCLATURE

RGB	Red Green Blue Value, a common way to show color according to the makeup of red, green, and blue with white being represented as 255, 255, 255 and black as 0, 0, 0.
Midrib	The central vein running along a leaf where smaller veins branch out from.

CHAPTER I

INTRODUCTION

The inspiration for this project arose in the Summer of 2017 when, as a part of Dr. Bruce Gooch's summer research team, I traveled to Ardmore, Oklahoma to visit the Noble Research Institute. I soon met with Dr. Stephen Webb and Dr. Charles Rohla and began talking about possible projects that could benefit from integrating with technology. The one that caught my eye was the analysis of pecan scab on leaves of infected pecan trees.

A Brief Overview of Pecans

Pecans are a native nut to North America, with over one thousand known varieties. [3] In 2014, the pecan crop in the U.S. totaled 264.2 million pounds, and at that time the value of the crop was $517 million, with the majority of production coming from Georgia, New Mexico, and Texas. [4]

The prominent history of pecans in Texas starts with Governor James Stephen Hogg in 1906 requesting that a pecan tree serves as his headstone, and that the nuts of the tree be distributed to make Texas a "Land of Pecan Trees.". In addition, the pecan tree became the official tree of Texas in 1919 in the context of the 36th Texas Legislature. [1] Texas has specifically seen a gigantic leap in pecan production, seeing their value of pecan production jumping from $44.8 million in 2013 to $107.8 million in 2014. [4]

The buyers of pecan mostly care about the nutmeat, which is why it's critical for farmers to know their shell-out percentage in order to have an idea of the value of their production. [4] Therefore, any way that farmers can increase their shell-out percentage by eliminating any possible hindrances to their crop, such as pests or diseases is critical.

Increasing the yield of orchards is becoming even more important as pecan exports overseas, mainly to Hong Kong and Vietnam, as exports were valued at $446.3 million in 2014 and are steadily increasing. [4] Despite the large increase in sales in the U.S. and overseas, the amount of research on pecans is surprisingly low. Dr. Rohla, who is one of the few experts in the world on pecans, emphasizes the importance on pecan research in order to increase supply in order to keep up with the rising demand.

Pecan Scab

One of the largest hindrances to pecan yield every year is the fungus pecan scab, also known as *Fusicladium effusum*. This fungus, commonly found in the American south, leaves dark spots on the leaves and nuts of affected trees and reduces respiration rates, nut number, nut weight, oil, moisture, and protein content. In addition, due to the effects on the leaves, the trees see a loss of not only photosynthesis, but also an increase in leaf drop. [5] Currently, the only remedy for the problem is fungicides.

The economic impact of pecan scab is great, such as in 2013 when there was an estimated 20-25-million-pound loss of pecans due to weather conditions that propagated good conditions for pecan scab. [2] The average lost due to pecan scab in orchards was around $22 million back in 1998 [6], with the number only rising due to the increase in the amount of strains of pecan scab that carries resistances to the fungicides.

The Problem

A problem that researchers face in the field is categorizing the infection level of pecan scab on the leaves of affected trees. Currently, the system used by most researchers is highly subjective, as ratings go off of their experience and judgement. As is clear, this causes a lack of standardization across the board for categorizing how infected trees are with pecan scab. A 20%

infected leaf to one researcher could be a 30% infected leaf to another. This inconsistency is extremely detrimental to trying to get consistent data in order to track the spread and infection levels of pecan scab across the south.

While there have been tools to help researchers standardize their readings, such as charts displaying infection levels developed by researchers from Georgia and Florida [7], these still rely on the judgement of the researcher. Though the charts do help, there's still variance as determining individual percentages remains difficult.

My Solution

In order to standardize the readings of infection levels across all researchers, the determination needs to shift from researchers who naturally possess varied opinions to an objectively consistent algorithm. I propose that the best method for determining the infection levels of pecan scab and making it consistent across all researchers is to have an algorithm that uses photo analysis to pinpoint the infection levels by comparing the leaf area to the dark pecan scab spots.

This algorithm will allow data on pecan scab to henceforth be completely objective and consistent, which will allow researchers to better determine the infection levels of pecan scab across the nation and know better when to tell farmers to spray fungicides and how much to spray. In addition, this consistent data will allow researchers to have a database to look at trends over years and match up the levels with weather data in order to forecast potential pecan scab outbreaks.

CHAPTER II

ALGORITHM DESIGN

Preliminary Decisions

One of the first decisions I had to make when starting the project was determining the programming language I was going to use. While there are numerous languages that are naturally helpful with included libraries to work with images and graphics, I ultimately decided to work in the coding language known as Java. Coding in Java allows me to do a few things, such as working in language that I'm already very familiar with, easy integration for working with images, and a simple transition to Android phones in the future as Java is used to program the mobile devices.

Another early decision was figuring out what type of data I wanted to be inputted into the algorithm to decipher, whether it would be a singular leaf, a multitude of leaves on a branch, or multiple leaves detached from a branch. This decision came down to the importance of accuracy for determining the infection levels of pecan scab. After looking at a multitude of images, I saw that trying to process multiple leaves on a branch would cause issues as the shadows created by the branch and leaves at different levels would create false positives for pecan scab. Therefore, I came to the conclusion that keeping the data input limited to either a single leaf or a small collection of leaves detached from a branch.

Data Collection

While initially this part of the project was going to be quite simple, gathering the required data took a different route than what I initially planned. I was working with Noble through the inception and progress of the project, but they were extremely limited in what they could share with me as I was an outside researcher working with data owned by them. Since the gathering of leaf data from Noble took more time than what I initially thought, I went ahead and used the program on pecan leaves from a database that I was able to find online. Surprisingly, these images worked perfectly and helped me fine tune and perfect the algorithm for the use of Noble researchers.

Algorithm Overview

The basis of the algorithm is decently simple, the algorithm needed to determine in the image what was the leaf, what was the pecan scab, and what was the background that the leaf was on. In order to make this process easier for the algorithm, I decided that it would be best for all the colors to be pretty distinguishable. To achieve this, the leaf needs to be on a white background and well lit, this would make the color difference between the background, leaf, and scab extreme. This difference in color allows for the algorithm to really only have to worry about three colors, white, green, and black. After determining what is leaf, scab, and background, the algorithm outputs the infection level of the leaf or average for the leaves in the image and outputs the information accordingly.

Algorithm Details

Since the overall goal of the algorithm is more or less explained, I'll start going more into the details. The algorithm first grabs the desired image that needs to be processed and sets up counters internally that keep track of how many pixels there are of each possible type, background, leaf, and scab. Then, the algorithm goes through every pixel found in the image, starting from the top left and moving to the right until it hits the last pixel of the image and then returns to the first pixel of the row below until the last pixel of the bottom row is reached. At every pixel, the color for the pixel is pulled. If the RGB value is above the threshold for white, which I initially set as 150 for red, 150 for green, and 150 for blue, the background counter increments. If the RGB value for the pixel was within the range I set for the leaf, which in this case was red being between 50 and 120, green being between 70 and 180, and blue being between 30 and 120, the leaf counter increments. I gathered these values by estimating from looking at some color found in different pecan leaves. Lastly, to determine the scab, if the pixel didn't fit any of the RGB ranges above, it would be marked as scab. Then, using the pixel counters, the algorithm calculates the values to output the total amount of pecan scab on the leaf. First, the algorithm adds together the values of the leaf counter and scab counter to get the total amount of the pixels that make up the leaf, then the amount of pecan scab pixels is divided by the total area of the leaf to give the infection percentage of the leaf or leaves. This information is then presented to the user to allow them to use in their research of pecan scab.

After writing the algorithm I decided to test it on some very basic color images that I made in Photoshop. These were meant to run very basic preliminary tests to see if the algorithm was working correctly and if my calculations were effective. So, I made images with three different levels of infection, 10% (Figure 1), 40% (Figure 2), and 80% (Figure 3). I then ran them through the algorithm and got results that almost exactly matched these percentages. This test showed promising results, and helped lay the groundwork for the tests that I would be performing later on the images of actual leaves. The test images are shown on the next pages.

Figure 1. 10% infected pecan scab test image

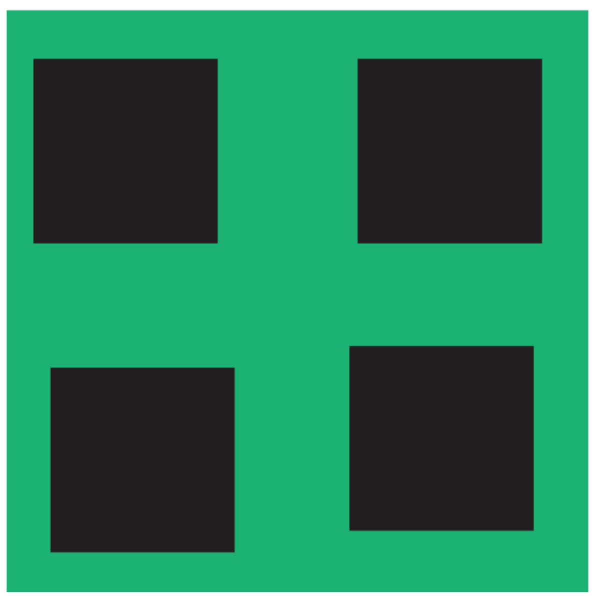

Figure 2. 40% infected pecan scab test image

Figure 3. 80% infected pecan scab test image

CHAPTER III

LEAF TESTING

Initial Tests on Leaves

After testing the algorithm on sample artificial images, I decided to test the algorithm on one of the leaf pictures (Figure 4) that I found online. What I found was that the infection levels were incredibly high for the amount of scab that was present on the leaf. The levels that were outputted were in the 80% range while I expected under 10%.

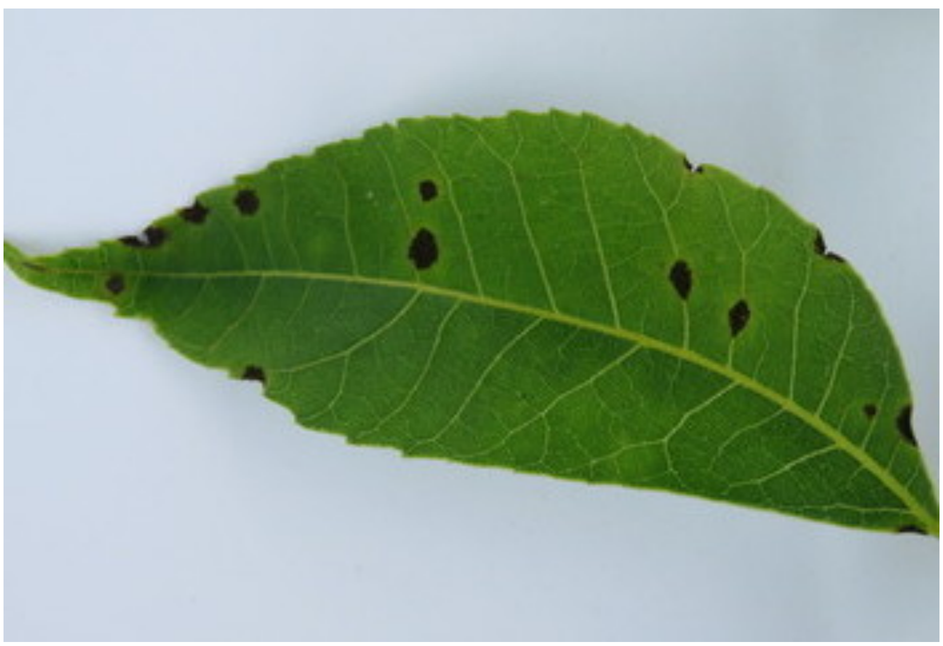

Figure 4. Pecan Scab Test Leaf 1 [8]

As you can see from the above leaf, the 80% infection rate is completely off. Therefore, I decided to see what my algorithm was counting as scab and what it was counting as leaf, so I wrote a process in the algorithm that draws a new image based on what type of pixel the algorithm reads. While the algorithm goes along, there'll be another image created to not only fix this problem, but also perfect the range for green and scab. With the new part of the algorithm written, I ran the program with the same ranges for green to see the drawing work (Figure 5). The picture that it produced is below.

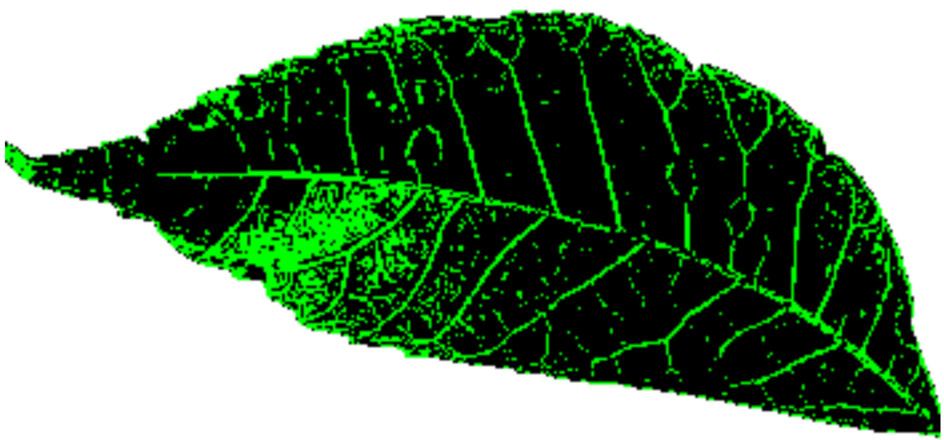

Figure 5. Algorithm Original Settings Drawn Image

First Leaf Filter Redesign

Seeing that the dark nature of the leaf contributed greatly to the false positives, I went to the pixel color device in Photoshop to look at the RGB values of the pixel to set up a new range for green. I quickly found that the artificial test images and the actual leaf varied greatly, as light plays a huge part in what RGB values get created. I therefore began the process of pinpointing a range that could be used to for the most part eliminate false positives created by dark regions of the leaf while still keeping the reading of the scab fairly accurate. After running the updated code, I received much more accurate results, producing the image below (Figure 6).

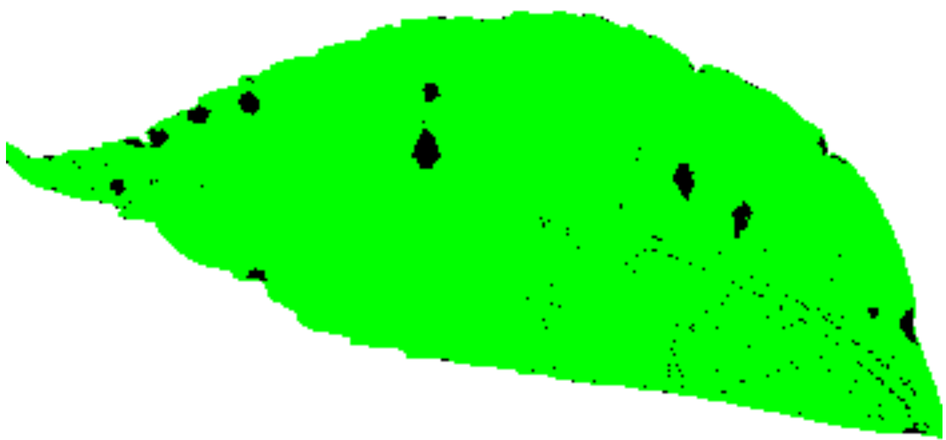

Figure 6. Algorithm Updated Setting Drawn Image.

As one can see, there still remains some false positives around the darker regions located near the base of the midrib, locations along the veins, and on some places of the edge. This image was the most accurate I could make the scab reading without affecting the area of the actual pecan scab on the leaf. The only way to reduce the number of false positives at this point, I thought, would be a better light source that eliminates the darker regions that are created by shadows from the midrib, veins, and edges of the leaf. I then decided to test my algorithm on other leaf samples to see the results.

Dealing with Light Colored Leaves

As I was testing the algorithm on leaf samples, I reached a sample where it was very well lit and had many bright greens make up the leaf (Figure 7). When I ran this leaf through the program, I was getting a high number of false positives for scab at the lightest locations for the green on the leaf (Figure 8). This error meant that there were thresholds of green that weren't registering in my filter for green, which warranted me having to revisit and test until I eliminated these errors. I reworked the filter by fixing my previous incorrect logic. Before, I was capping the max values for red, green, and blue, thereby limiting the brightness of the values. I had completely forgotten that the algorithm would have already determined all the background spaces, meaning that I would need no upper bound on the values for red, green, and blue, which meant that I could catch any shade of green that was well lit with no problem. After testing, I set the background filter to have the RGB values to catch the values that were all above 110, and the leaf filter to catch values of those pixels where red was above 15, green above 65, and blue above 0 (Figure 9). Below are the images showing my results.

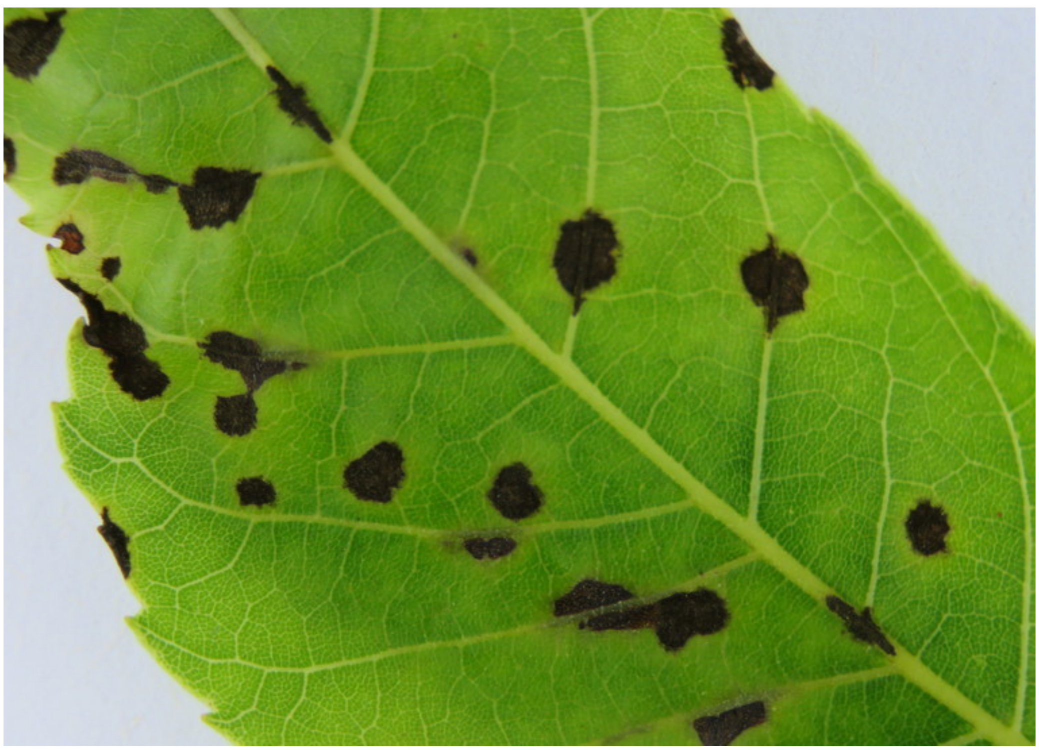

Figure 7. Pecan Scab Test Leaf 2 [8]

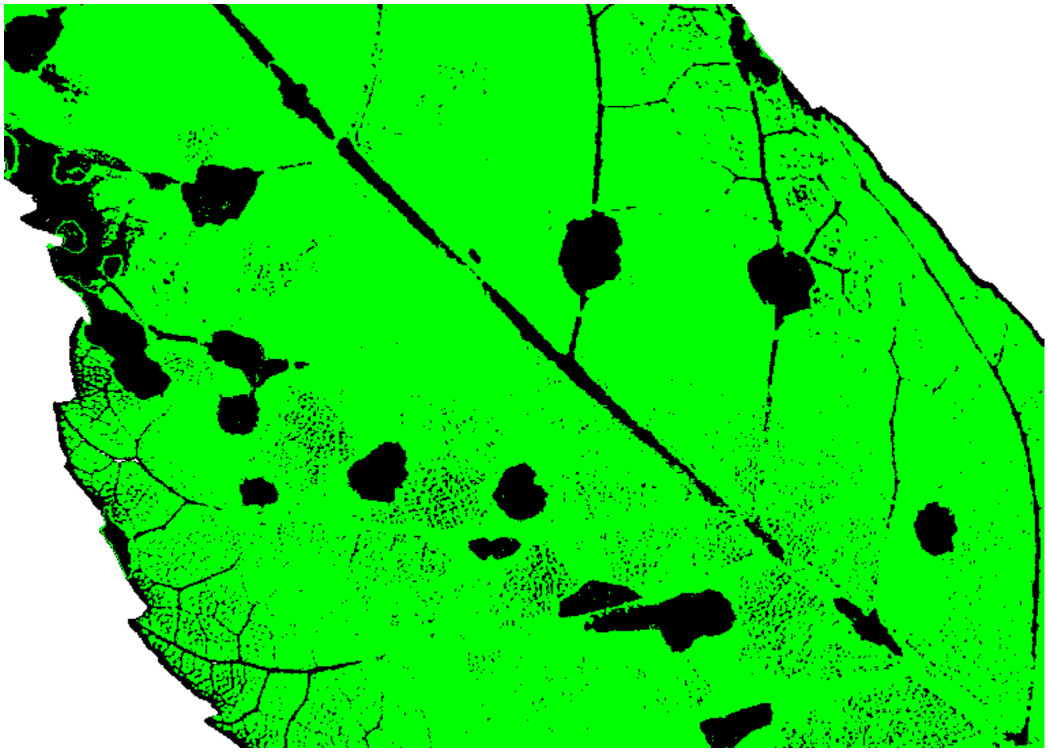

Figure 8. Algorithm Drawn Image Leaf 2 with Leaf Filter Values Capped

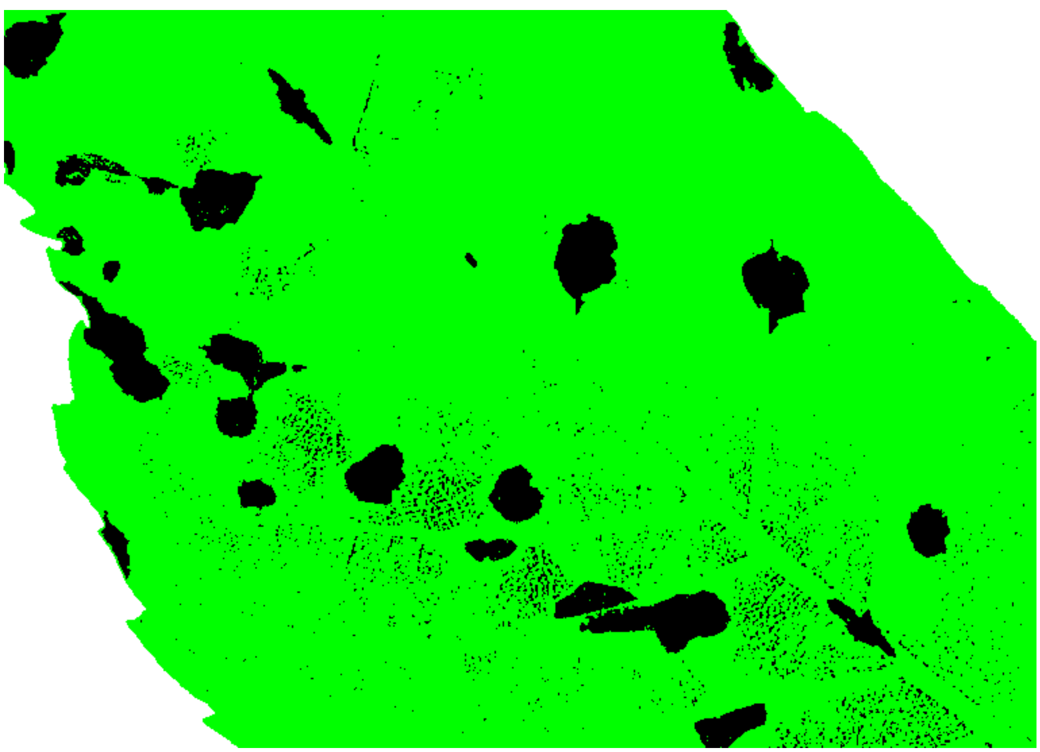

Figure 9. Algorithm Drawn Image Leaf 2 with Leaf Filter Values Un-Capped

Fixing False Positives

As you can see above, I was able to significantly reduce the number of false positives from bright leaf spots, but I was still left with these black dots that registered as scab. I originally thought that these spots were a result of having my catch for green being too low, but after testing a multitude of different number combinations, there was nothing that I could do to remove these spots, even when I set the values to make everything on the leaf green. I then decided to try testing numbers in the negative to see if somehow my color conversion was somehow putting some values below 0, and ended up being the answer to the leaf spots registering as scab. So, I set a new value that set all the values remaining for red, green, and blue, that were above 0 to be scab, and anything else, in this case the leaf values that somehow had these negative values, to be leaf (Figure 10). The image below shows how much this cleaned up the false positives.

Figure 10. Algorithm Drawn Image Leaf 2 with Random Spots Eliminated

With this final clean up, I finally produced an algorithm that I was satisfied with to determine what was scab on a pecan leaf. The algorithm showed that leaf 2 was 6.95% infected with pecan scab. While the algorithm's performance under low light conditions isn't the best, when the leaf is properly lit up with a good light source, it does a fantastic job precisely determining the infection levels of pecan leaves with pecan scab. Below are some more leaves that I ran the algorithm on, with their original image and algorithm drawn image. The infection results for the leaves are as follows, leaf 3 (Figure 11) is 7.15% infected (Figure 12), leaf 4 (Figure 13) is 6.17% infected (Figure 14), and leaf 5 (Figure 15) is 7.27% infected (Figure 16).

Figure 11. Pecan Scab Test Leaf 3 [8]

Figure 12. Algorithm Drawn Leaf 3, 7.15% infected

Figure 13. Pecan Scab Test Leaf 4

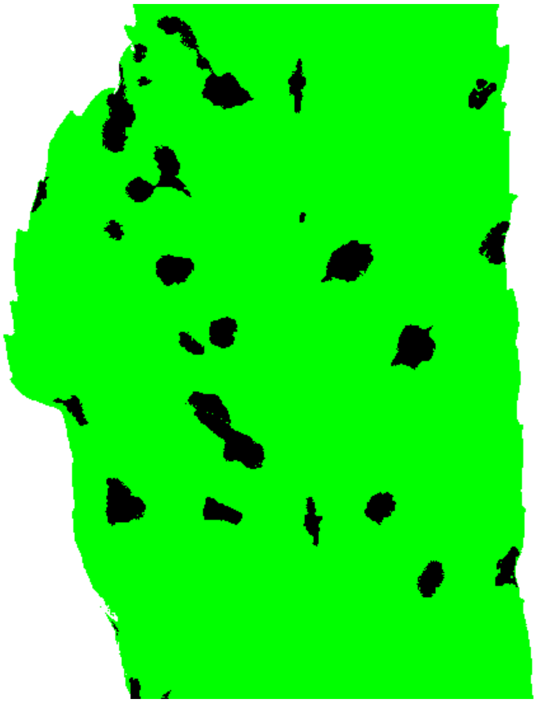

Figure 14. Algorithm Drawn Leaf 4, 6.17% infected

Figure 15. Pecan Scab Test Leaf 5 [8]

Figure 16. Algorithm Drawn Leaf 5, 7.27% infected

CHAPTER IV

CONCLUSION

With pecan scab becoming an ever-increasing threat to the production of pecans in the Southeastern United States, understanding and fighting this disease is critical for allowing the supply of pecan nuts to keep up with the incredible amount of demand. As researchers continue to find solutions to curb and eliminate this fungus from infecting orchards and affecting farmers across the Southeast, having standardized data for infection levels is vital for tracking this disease over time, comparing results from reduction attempts, and eliminating the innate subjective nature of judging infection levels by eye. This program gives the solution to researchers in the field and allows them to precisely determine the infection levels of their leaves to better understand pecan scab. After trial and error, the algorithm produces extremely accurate results as long as the tester makes sure to control the environment and produce a satisfactory light source to shine on the leaf. While there are still some false positives produced when the leaf isn't properly illuminated, these false positives aren't a result of the algorithm itself but of the light on the leaf. The implications of this paper shouldn't be limited to pecan scab, but the photographic analyzation of leaves can be adapted and molded to all types of possible diseases for other types of plants. If researchers all over can better determine infection levels of diseases on plants, research can become more accurate and help farmers all across the globe.

REFERENCES

[1] "FAQs." Texas Pecan Growers Association, Texas Pecan Growers Association, 2002, www.tpga.org/faqs.php.

[2] Thompson, Clint. "Pecan Crop Hurt by Summer Rainfall." College of Agricultural and Environmental Sciences, University of Georgia, 8 Sept.

[3] "Pecan Varities." Menu Pecan Varieties There Are, 2010, www.texaspecans.org/varieties.html.

[4] Marzolo, Gina. "Pecans." Ag Marketing Resource Center, Aug. 2015, www.agmrc.org/commodities-products/nuts/pecans/.

[5] Bock, Clive. "Fusicladium Effusum (Pecan Scab)." Fusicladium Effusum (Pecan Scab), www.cabi.org/isc/datasheet/13719.

[6] "USDA ARS Online Magazine Vol. 46, No. 8." AgResearch Mag, Aug. 1998, agresearchmag.ars.usda.gov/1998/aug/scab/.

[7] Yadav, et al. "Development and Validation of Standard Area Diagrams to Aid Assessment of Pecan Scab Symptoms on Fruit." Plant Pathology, 2013.

[8] Jonas Janner Hamann, Universidade Federal de Santa Maria (UFSM), Bugwood.org