DETECTION OF DISTRESS AND DISEASE IN DECIDUOUS TREES UTILIZING REMOTE SENSING

By

JACOB MORRIS

MASTER OF SCIENCE
GEOGRAPHIC INFORMATION SYSTEMS TECHNOLOGY
FINAL PROJECT

THE UNIVERSITY OF ARIZONA

2022

I dedicate this project to Michael Vay Morris.

ACKNOWLEDGMENTS

I thank my partner and parents for the encouragement and reminders over the years to ensure that I achieved my personal ambitions and goals.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

GSD	Ground sample distance
KNN	K-Nearest Neighbor
NAIP	National Agriculture Imagery Program
NIR	Near Inferred light spectrum
SVM	Support Vector Machine
UGRC	Utah Geospatial Resource Center

ABSTRACT

Locating and identification of plant stress and diseases plays a major role in plant conservation and human safety concerns relating to falling hazards and reduction in fire blocks between structures in medium sized population centers. Overall flora heath can be indicated by visual observations of the chlorophyll and other pigments in the leaves. As outside interference with the plants ability to naturally produce the required nutrients, such as environmental and pathological interference, the visible pigmentation change. In this study, pigment variation is evaluated and analyzed by machine learning methods including image classification for the evaluation of health in deciduous trees. By utilizing multispectral imagery this study compares wavelength values for identified affected individuals showing visual symptoms to located other affected individuals both showing symptomatic and non-symptomatic individuals. Data analysis was conducted utilizing a trained supervised classification, support vector machines and K nearest neighbor method to determine which methods was most precise in identifying affected pixels for fast-tracking management evaluations for resource managers. The overall classification accuracy of targeted, healthy, fields, and urban was relatively good, with kappa values ranging from 0.66 to 0.75 and overall accuracy ranging from 70%to 83%. Support vector machines accuracy of 82.13% with a kappa coefficient of 0.74 at a 750 point accuracy assessment making it the best method of the two for detections of symptomatic and asymptomatic individuals.

Keywords: distressed detection, disease detection, machine learning, multispectral imaging

ETHIC STATEMENT

Present day Geospatial Information Systems (GIS) has simplified everyday tasks such as providing quick access to maps. With this increase of access, and the inherent belief that published maps are truthful, the ability to violate this belief is easily done by distortion either intentionally or unintentionally. Furthermore, with access to a plethora source of information, this can create unintentional violations to privacy, unjust racial targeting, and intellectual property theft (plagiarism). This study is no exception to these potential ethical concerns. This project will following the ethical principle of always treating other with respect, by considering the impact of the study's actions on others. To achieve this, the researcher needs to evaluate our obligations to others (colleagues/professionals, institutions, and society), to provide the best work possible (full use of education/skills, deliver quality work) and to provide honest representation. A seven-step evaluation process is utilized to determine if an action is ethical or unethical and should be omitted.

During this study there were many opportunities that could be identified as potential unethical situations. Such situations include not providing proper credit owed to other researchers work for providing direction in the study, not listing out all the processes involved for others to reproduce, or misrepresentation of data including the deletion of outlier data. As part of any scientific process the ability to properly present and allow for peers to reproduce is key. If at any point the researcher fails to properly represent the studies' process/results, it invalidates the efforts of themselves and their fellow colleagues. The biggest ethical concern for this study is the overall impact, interpretation and usefulness to the intended audience and the public. This study is intended to assist and inform resource managers of an urban center of specimens that

may need evaluation. However by providing inaccurate workflow data and results this could lead to excessive cost to society, your client and to your peers in GIS.

CHAPTER 1
INTRODUCTION

With much of the population in the United States moving to urban centers (an estimated 81% of the total population now live in or near cities), urban forests are becoming more important than ever. By creating a network of dynamic ecosystems, they filter air and water, assist in mitigating storm effects, conserve energy, and provide habitats and shade [1]. Over much of northern Utah, urban centers utilize several species of deciduous trees that are native to the area. Two of the most common are form the Acer genus. This genus is well suited for a variety of purposes like mountain stream protection *(Acer glabrum)*, ability to tolerate certain level of drought and poor soil composition *(Acer grandidentatum) and* providing wildlife with ample habitats [2]. With local and federal resource managers playing a larger role in the preservation of natural aesthetics, recreational uses, and public safety concerns, this has become a more challenging task than ever before.

Logan City, Utah has been awarded “Tree City USA” for over 30 years and continues to increase its urban forest development by 100 to 200 new trees annually [3]. With this continual growth in the urban forest and a continual population growth of 8.7% from 2010 (48,174) to 2020 (52,778) [4], management and public safety concerns start to surface. With influences both external and internal, many of these deciduous trees can become compromised and create unknown safety hazards.

Like many species of flora weather, climate and changes over time can create unfavorable conditions resulting in distress of the organism [5]. With some tolerance to drought which is the most common climate concern for Northern Utah, *Acer grandidentaum* is preferred in much of the urban centers in valleys [2]. With common

“climate-stress” diseases, an external pathogen is not needed to create physiological impairment (such as scorched leaf disease). External stressors such as pathogens are generally incapable of mortality in health specimens, but they are more favorable in specimens experiencing climate stress [5].

GIS has aided in the development of early and current detection of disease by utilizing various imaging techniques (such as hyperspectral imagery and machine vision) and machine learning [6]. This is true especially in trees and other agricultural crops that shows unique foliage discoloration.

Jaafar Abdulridha et al. (2019) utilizes several of these techniques in the detection of citrus canker disease. In this study the utilization of both field collection and remote sensing was implemented to locate and verify the condition of the suspected specimens. Unmanned Arial Vehicles (UAV) were implemented to take imagery 30 meters above the canopy. This imagery was then evaluated in various ways including reflectance observations, vegetation indices evaluating chlorophyll reduction, spectral classification analysis, supervised machine learning, and k-nearest neighbor.

To evaluate the effectiveness indoor imaging techniques were used to create the spectral signatures of different stages in a controlled environment. This was important to assist them in understanding how each stage’s reflectance changed in development over the course of the disease.

M.L. Perez-Bueno et al. (2019) conducted a study that utilized both field collection and aerial imagery. Field samples were collected to verify infected trees by placing a sick in the ground near a suspected infection site and collecting samples for examination in a laboratory. This was completed to verify the results of the remotely

sensed analysis. Regarding the remote sensing aspect of this study, they utilized both Normalized Difference Vegetation Index (NDVI) and Normalized Canopy Temperature (NCT). In addition to these methods, logistic regression analysis was used to assist in the detection of potentially affected specimens.

Xiaoling Deng et.al. (2020) utilizes remote sensing techniques to study imagery that was obtained using a UAV that was mounted with a GPS enabled camera. Once the images were gathered, they were then processed by utilizing a deep learning framework, to identify individual subjects that were affected with the disease within the study region.

Thomas et al. (2018) explores how non-invasive optical sensors are available which assess solar reflectance properties in plants by evaluating different regions of the electromagnetic spectrum. By evaluating these regions this study explores precision in the detection of plant stress in agricultural crops. Imagery such as images from the National Agriculture Imagery Program (NAIP) which captures images during peak growing season would be ideal for studies like the Thomas et al. study [11].

This study will determine if using remote sensing techniques in combination with machine learning can improve the ability for city resource managers to identify potential hazardous trees within the study area. This process will evaluate multispectral imagery acquired from NAIP for the Logan City area by utilizing reflectance wavelengths of canopy cover to identify potential symptomatic or asymptomatic individuals through leaf reflectance changes in deciduous trees.

CHAPTER 2
METHODS AND DATA

Workflow

Initially the collected data was for the entire Cache County, Utah area. For this study and for simplification purposes the project only looks at the City of Logan NAIP imagery, which is often collected with just red, green, and blue bands. However, some flights include a near infrared band (NIR) which is very useful for quantifying vegetation cover and health [12]. NIR is a portion of the light spectrum that is just outside of the human eyes ability to process and have a wavelength spanning approximately 0.7 to 0.14 micrometers. To evaluate the desired study area, a portion of the NAIP raster was clipped using the raster clipping tool on the NAIP 2021 rasters leaving only data within the Logan City limits.

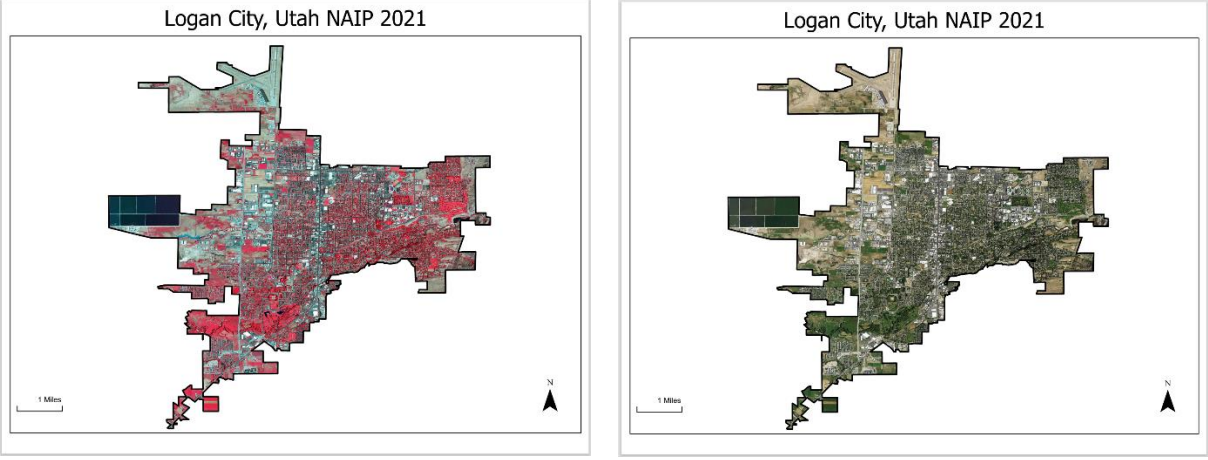


Figure 2-1. Both natural and near infrared overlays for the City of Logan Study area.

This was done to both help in the evaluation process and improve the overall speed of the tool processing. The following figure shows the overall procedure in this study in the evaluation of the area of interest.

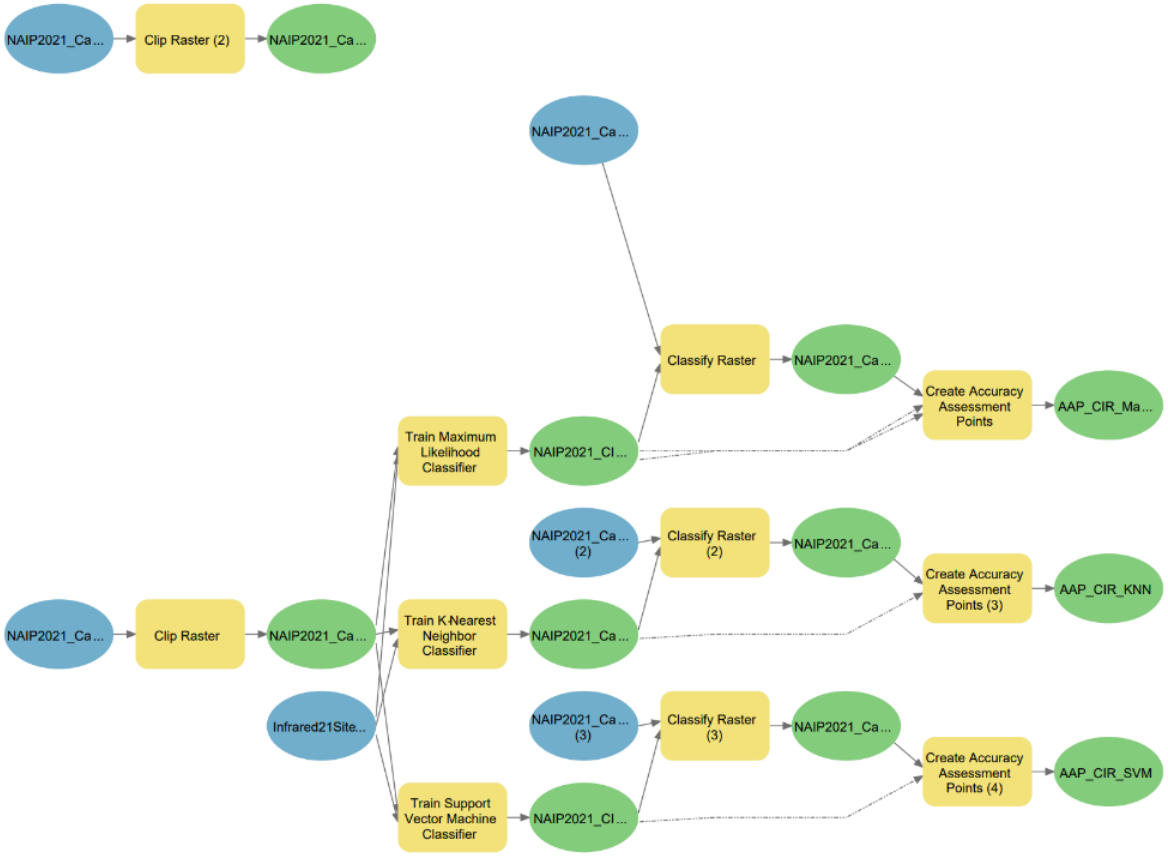


Figure 2-2. Workflow process for evaluating and identifying affected individuals

The techniques

Detection

The detection process of this project started off by visual observations of the NIR imagery. NIR imagery was utilized for this process instead of natural color (RBG) imagery after determining which type of imagery was most suited for evaluation of vegetation health.

Figure 2-3. Locating targeted training points (light pink centered tree) in the study area

After locating potential symptomatic individuals, the city forester was contacted to see if there were records of symptomatic individuals to compare to see how accurate the visual observations were. The city forester noted that there was not a current record of declining trees, and that larger cities may have records for their respected area. The start of the evaluation process for the rasters was the establishment of 120 training sites (30 Targeted, 30 Healthy, 30 Fields and 30 for Urban). This was necessary so the classifiers could be trained in accordance with visual observations.

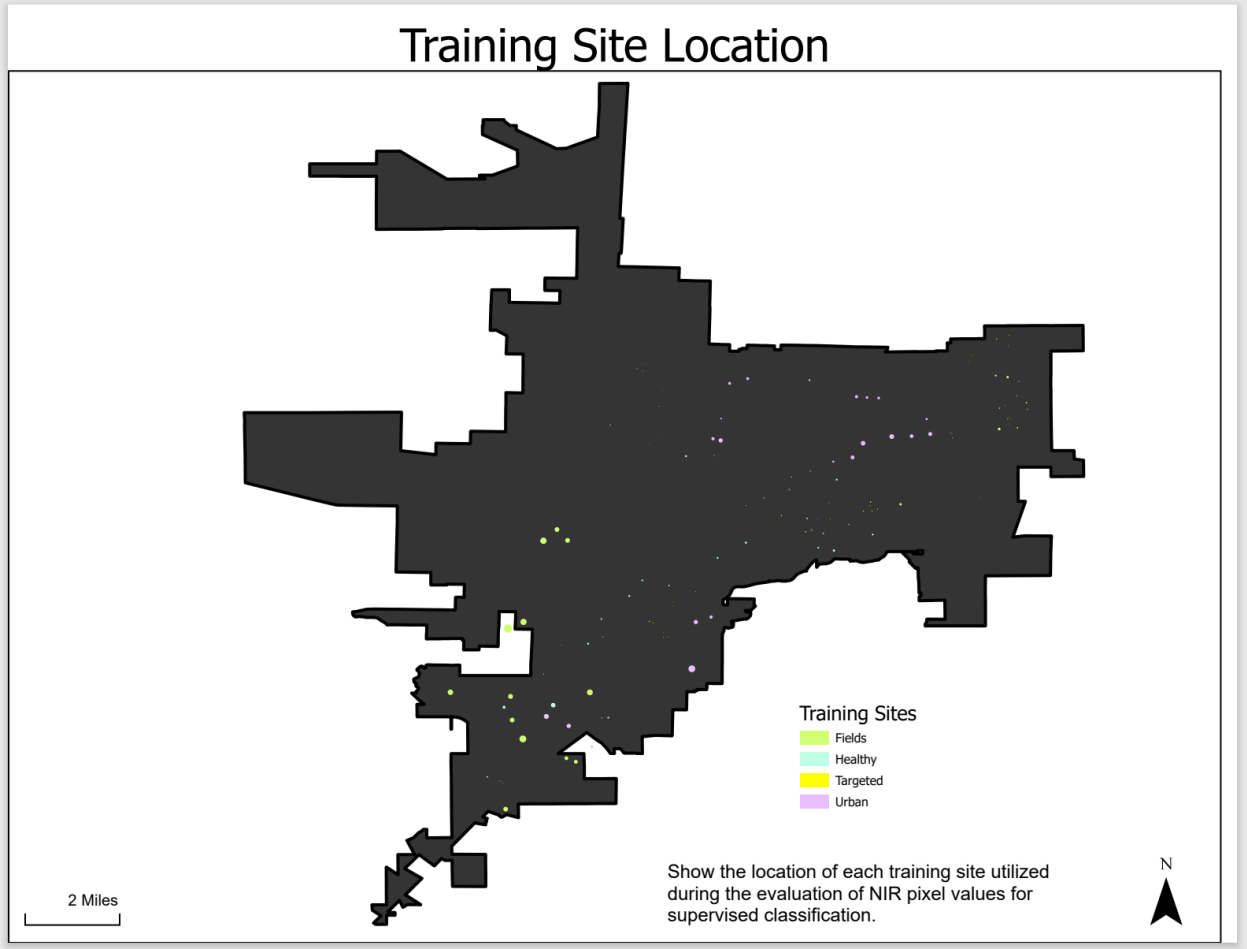


Figure 2-4. Shows the location of each training site utilized in the training process of the classifiers.

Analysis

With the establishment of the training sites, various analytical processes were conducted and evaluated to identify other affected specimens in the study area. The software ArcGIS Pro (version 2.9.0, Esri, Redlands, CA, U.S.A.) was used for the analysis and observational process. Two types of supervised classification analyses were preformed to evaluate which method would identify most accurately. The first analysis that was preformed was a Support vector machine (SVM). SVM classifier is a supervised classification method that is well suited for segmented rasters but can also handle standard imagery. In addition, SVM has advantages over other methods such as

requiring fewer samples, not requiring the samples to be normally distributed, and is less susceptible to noise [17]. The second analysis was the K-Nearest Neighbor (KNN). KNN is a nonparametric classification method that classifies a pixel or segment by a plurality vote of its neighbors. K is the defined number of neighbors used in voting [18]. Figure 2-5 shows the classification of pixels based on the SVM method and Figure 2-6 shows the classification of pixels based on the KNN method.

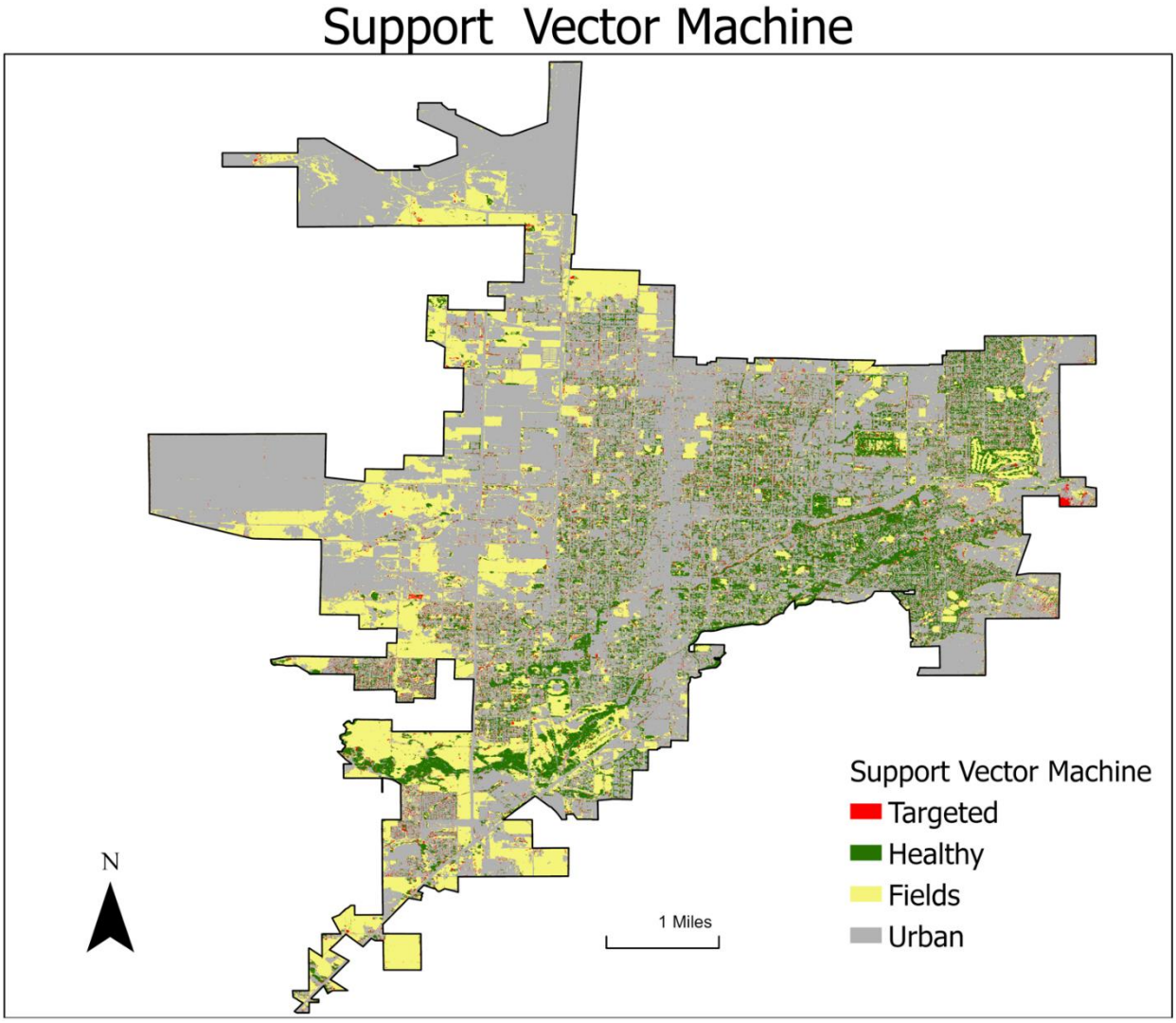


Figure 2-5. Classified pixels determination by using the Support Vector Machine classification method.

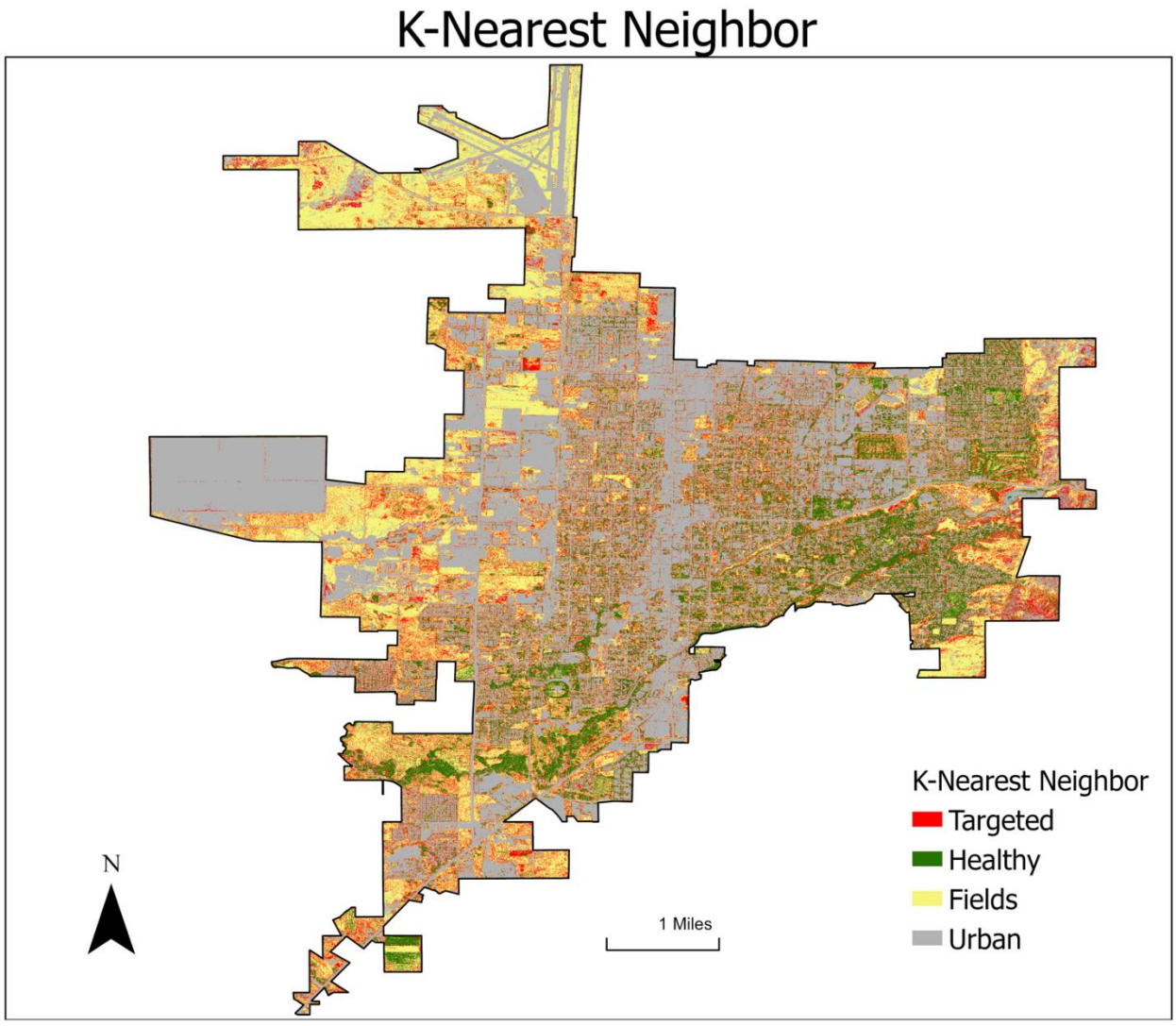


Figure 2-6. Classified pixels determination by using the K-Nearest Neighbor classification method.

Microsoft Excel was used for statistical evaluation of each of the classification methods. Accuracy assessment points were created in multiple levels (500, 750, and 1,000 points). A confusion matrix for each method was done and all classes were evaluated for overall accuracy, users' accuracy, producers' accuracy, and the kappa coefficient. Overall accuracy is the proportion of correctly predicted values in relationship to the total amount of tested values. This however can lead to

shortsightedness and hides details. Users' accuracy evaluates how accurate the reclassified map is compared to the ground truth. Producers' accuracy is how accurate a map is from the authors point of view. Kappa coefficient measures the agreement of classification and the truth. Precision, also known as user accuracy, described the relationship between true positives (or correctly identified values) and the total number of true positives and false positives (incorrectly identified value deemed correct). Recall, also known as producer accuracy, is defined as the relation of true positives to the total number of true positives and false negatives (saying a pixel is incorrect when they are in fact correct). [13,19, 20]

The justification

Of the available data pertaining to this study, the National Agriculture Imagery Program (NAIP) was selected as the best fitting imagery. This imagery was selected because of the minimum spatial resolution requirements as of 2011 of ½ meter or less from ground sample distance (GSD), and that all imagery acquired is collected during peak growing season [11]. The infrared band was determined as the most efficient way to observe affected individuals due to its ability to detect how an object responds to Near infrared (NIR) light (i.e., absorbs, transmits, or reflects) which can reveal land cover conditions that are undetectable on natural color imagery. For example, it can identify stressed vegetation, moist areas in fields, and identify plants (e.g. differentiate between hardwoods and conifers) [13].

Supervised classification at pixel level in an image has the ability to classify pixels into characteristics and interpret them to define by the annotations of the training classes to every pixel [15]. This leads into the first method used in this study: the SVM. This method was chosen due to its risk minimization and its reliability in finding class

boundaries with a low number of training data. SVM has also been successfully applied in the identification of foliar diseases in agricultural crops [15].

The second method utilized in this study is the KNN. KNN was selected after reviewing a study from Jaafar Abdulrahim et al. (2019). In this study they evaluated several different methods in the identification of Citrus Canker Disease which show cases with similar symptoms to the factors being evaluated in this study [7]. Of the various methods reviewed, the current available data sources and the current software access available to this study were main factors for the selecting both SVM and KNN.

Data

All the NAIP data obtained for this study was collected form the Utah Geospatial Resource Center (UGRC) which created a mosaic imagery of images from the USGS database. Figure 2-7 shows the study area for 2021 NAIP with the near infrared band, and Figure 2-8 represents the NAIP – Natural color imagery for the same periods.

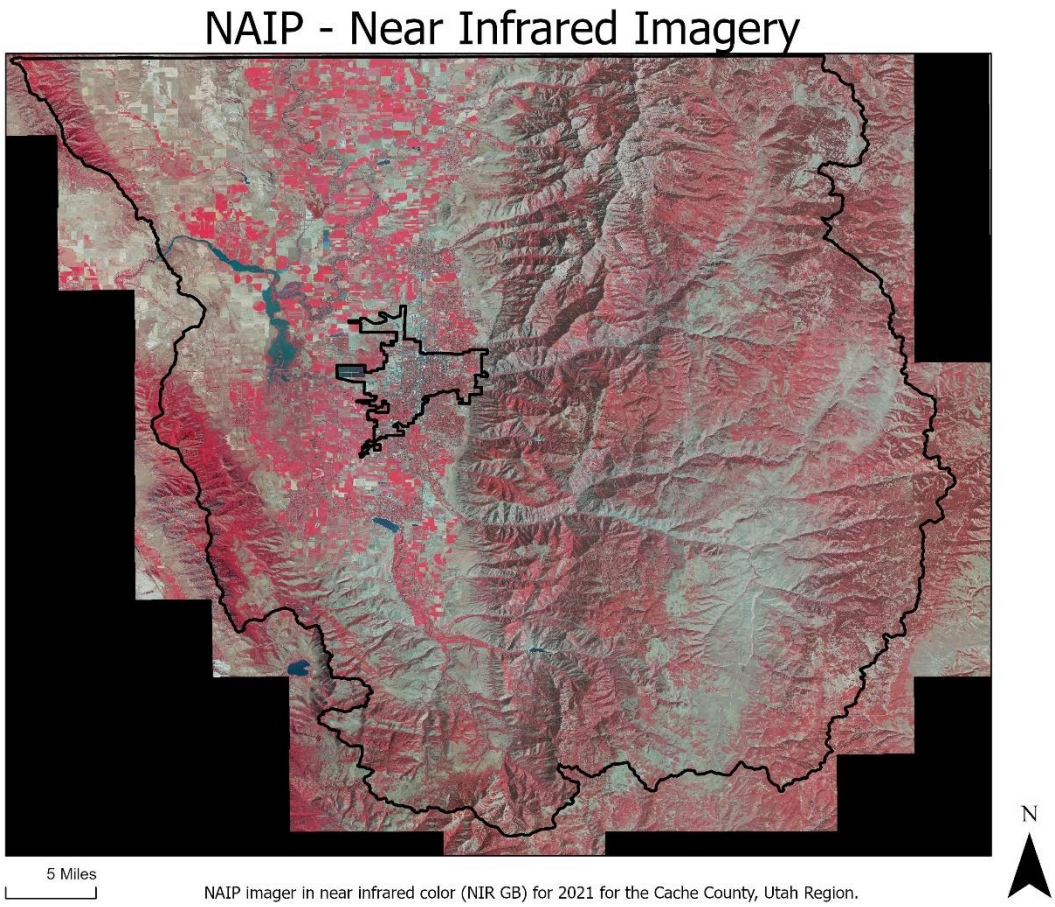


Figure 2-7. NAIP imagery with Near Infrared for 2021.

Table 2-1. Metadata for 2021 NAIP NIR imagery. This data set shows the recorded near infrared reflectance of vegetation for Cache County, Utah aiding in plant health indication.

Name	.6 Meter Color Infrared Digital Orthophotography form 2021 NAIP (county mosaics)	
Year	2021	
Author/ Owner	UGRC	
URL	https://gis.utah.gov/data/aerial-photography/naip/#NAIP2021	
Coordinate System	UTM zone 12	
	Datum/ellipsoid	WGS 84
	Horizontal Accuracy	+/- 4 meters to true ground
	Resolution	0.6 meter

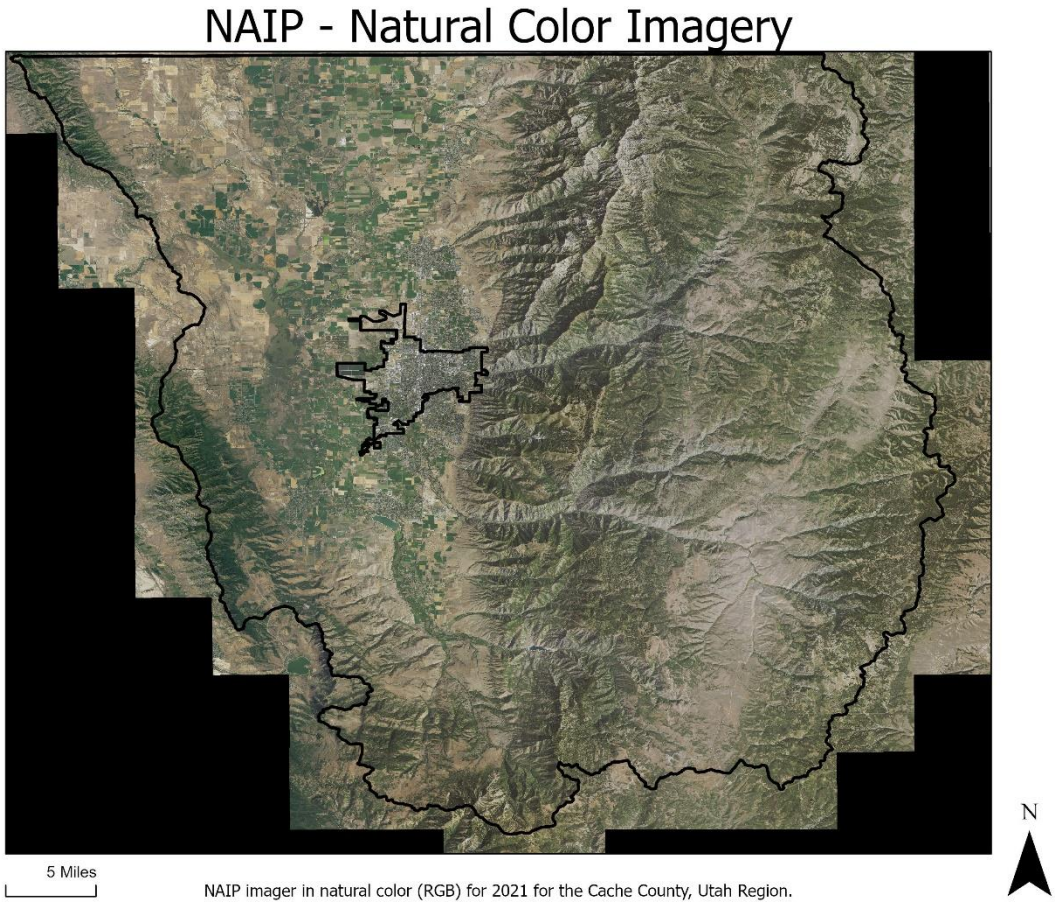


Figure 2-8. NAIP imagery in Natural Color (RGB) for 2021.

Table 2-2. Metadata for 2021 NAIP natural color imagery. This data set shows the a high resolution of natural colored imagery of Cache County, Utah.

Name	.6 Meter Natural Color Digital Orthophotography form 2021 NAIP (county mosaics)	
Year	2021	
Author/ Owner	UGRC	
URL	https://gis.utah.gov/data/aerial-photography/naip/#NAIP2021	
Coordinate System	UTM zone 12	
	Datum/ellipsoid	WGS 84
	Horizontal Accuracy	+/- 4 meters to true ground
	Resolution	0.6 meter

CHAPTER 3
RESULTS

To determine which of the two types of machine learning was most effective in locating symptomatic and asymptomatic individuals within the study area, randomly assigned assessment points were established and evaluated to determine accurate ground truth values. With these evaluated assessment points the creation of a confusion matrix made it possible to devise various statical appraisal for each model. Table 3-1 shows the layout of interpretation for the confusion matrix that was utilized in this study.

Table 3-1. Confusion matrix for predicted and actual values of accuracy assessment

		Actual			
		Affected	Healthy	Fields/Lawns	Urban
Predicted	Affected	TP	FP	FP	FP
	Healthy	FP	TP	FP	FP
	Fields/Lawns	FP	FP	TP	FP
	Urban	FP	FP	FP	TP

TP = true points or points that were accurately predicted, FP = false points or points that were not accurately predicted.

Determination of the accuracy that each model had in identifying symptomatic and asymptomatic individuals was evaluated by obtaining the percentage of correctly predicted points form the total points evaluated (Accuracy), type one errors (user's accuracy, u-accuracy), type two errors (producer's accuracy, p-accuracy), and kappa statistics. Kappa coefficient is a measure of how closely the instances classified by the machine learning classifier matched for the accuracy of a random classifier. Initially 100 random points was used to determine the overall ground truth versus classified values to see if more accurate training samples were needed. Table 3-2 shows the statistical data derived from this initial evaluation process.

Table 3-2. Confusion matrix for KNN initial 100 assessment point.

KNN 100 Assessment Points

	Affected	Healthy	Fields/Lawns	Urban	U-Total	U-Accuracy	Kappa
Affected	11	0	3	0	14	0.78571	
Healthy	0	21	1	0	22	0.95455	
Fields/Lawns	13	2	21	2	38	0.55263	
Urban	1	2	0	23	26	0.88462	
P-Total	25	25	25	25	100		
P-Accuracy	0.44000	0.84000	0.84000	0.92000		0.76000	
Kappa							0.68000

Table 3-3. Confusion matrix for SVM initial 100 assessment point.

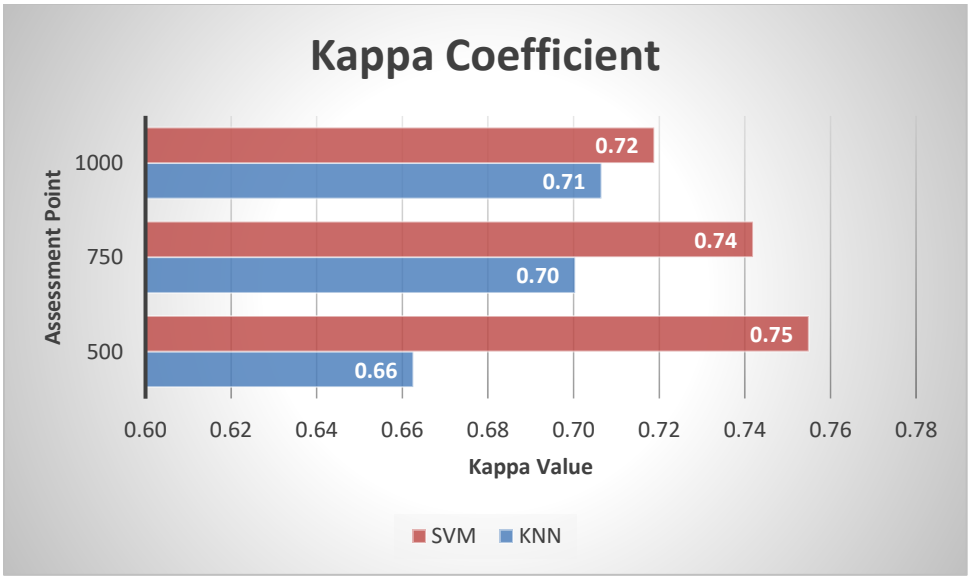
SVM 100 Assessment Points

	Affected	Healthy	Fields/Lawns	Urban	U-Total	U-Accuracy	Kappa
Affected	10	0	3	0	13	0.76923	
Healthy	0	17	2	0	19	0.89474	
Fields/Lawns	13	7	18	1	39	0.46154	
Urban	2	1	2	24	29	0.82759	
P-Total	25	25	25	25	100		
P-Accuracy	0.40000	0.68000	0.72000	0.96000		0.69000	
Kappa							0.58667

Based on this initial evaluation the observed accuracy of KNN was 76% whereas the SVM was 69%, and a kappa value of 0.68000 and 0.58667 respectively. Looking further into this evaluation, the producer's accuracy or P-accuracy tells us the likelihood that a value, in a given class, was classified correctly. From this you can determine the portion of values that belong to one classification but were predicted to be in a different classification otherwise known as omission errors. In this initial evaluation for both methods, the Affected category had a low P-accuracy of around 40% or having a high omission rate of approximately 60%. User's accuracy or U-accuracy looks at how valid

the predictions are, presented by ratio. Utilizing the U-accuracy you can determine the values that were predicted to be in a class but do not belong to that class also referred to commission errors. In this initial evaluation, for both methods, results show that most of the categories (except Fields/Lawns) have a U-accuracy greater than 75%. This tells us that most of the predicted values ended up correctly classified. In contrast this means that these categories have low commission errors. Individually each calculation shows you useful information but in order to determine if the test training sites are sufficient for this study, evaluations of all calculations together overall accuracy, kappa, user's accuracy, producer's accuracy and both types of errors needs to occur to ensure the reliability of the training process [21].

Due to low accuracy and kappa evaluating at moderate agreement, new training sites were established and new KNN and SVM classifications were compiled, and appraisal of these new classifications was done by utilizing 500, 750 and 1,000 assessment points. Figure 3-1 shows the statistical evaluation for both models at each respected assessment point.



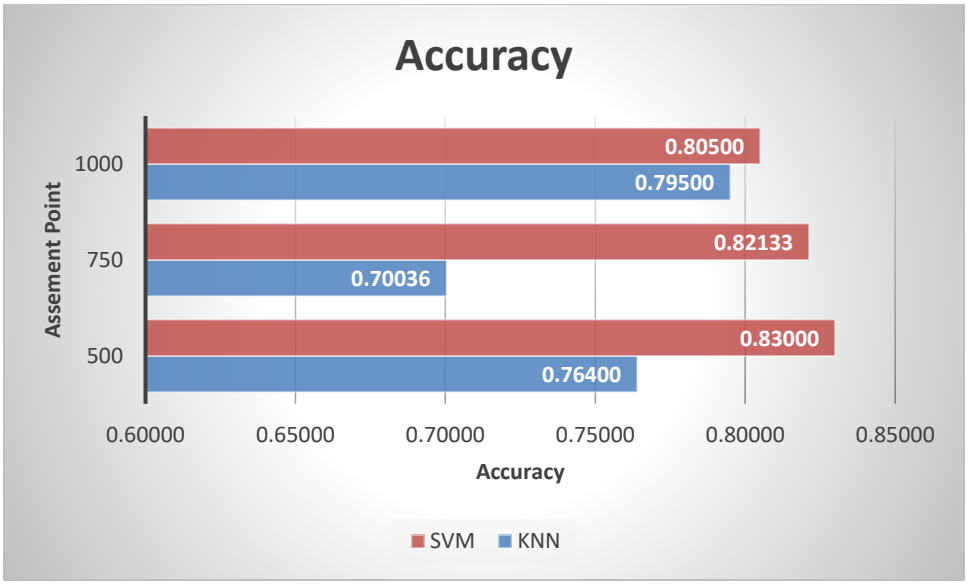


Figure 3-1. Statical evaluation for accuracy and kappa coefficient for both models at each level of assessment.

At the 500-point level there are notable changes in the accuracy of the data. The observed accuracy for KNN had no discerned change, whereas SVM had a significant increase in accuracy from 69% in the initial 100-point assessment to 83% in the 500-point assessment. A similar trend was present in the kappa coefficient with a slight decrease in the KNN kappa of about 0.02 and an increase in SVM kappa of about 0.14.

At the 750-point level the accuracy of the KNN was 70.03% with a kappa coefficient of 0.70 and the SVM had an accuracy of 82.13% with a kappa coefficient of 0.74. At the 1,000-point level the accuracy of the KNN was 79.50% with a kappa coefficient of 0.71 and the SVM had an accuracy of 80.50% with a kappa coefficient of 0.72.

The overall classification accuracy of targeted, healthy, fields, and urban was relatively good, with kappa values ranging from 0.66 to 0.75 and overall accuracy ranging from 70% to 83%. Both models were able to accurately determine individual/clusters of interest, with SVM maintaining a higher accuracy than KNN. With

higher accuracy of detection, it would be effective for resource management to utilize SVM classification in similar project.

CHAPTER 4
CONCLUSION

The aim of this study was to evaluate two methods of supervised classification to determine which method was most accurate in identifying symptomatic and asymptomatic deciduous trees in Logan City, Utah as an aide for resource management. Evaluating SVM and KNN output rasters by visual observations and statistical analysis shows that the SVM method was more efficient in identifying individuals with symptomatic and asymptomatic traits with higher overall accuracy and Kappa values. However, it is important to remember that SVM accuracy did have a slight decrease as the assessment points increased. This is believed to be due to similar values found in portions of fields, parks and lawns that were experiencing similar chlorophyll degeneration.

This particular study encountered a few limitations that restricted the overall application. The first finding adequate imagery during the peak growing season, that had minimal atmospheric distortion. Our second limitation during the initial application for classification was the rate of chlorophyll degradation, in both fields and lawns. This data showed similar NIR reflectance as symptomatic individuals. In addition, some technological limitations provided challenges for the study. These limitations included software restrictions, current programing availability within ArcGIS Pro, and hardware limitations such as sensor capturing.

This study provides an accurate way for management teams to quickly identify potential hazards and clusters of symptomatic individuals to maintain control. To further future studies, physical collection and study of specimen samples could be conducted to identify causes of the vegetation pigment variation within symptomatic and

asymptomatic individuals. By obtaining this information the study can start to create a targeting ability for a particular type of stressor within their region of interest. In addition to cause identification, generating a vegetation reflectance database and incorporating it within the study could potentially lead to species specific evaluations in order to track patterns and overall health. This study could also be utilized with modifications to aide in the monitoring of land within an urban center for various applications, such as erosion evaluation, native and invasive species tracking and more.

LIST OF REFERENCES

1. *Urban forests*. (2015, September 29). US Forest Service. https://www.fs.usda.gov/managing-land/urban-forests

2. *Urban Forestry Utah Native Trees*. Slc.Gov. Retrieved April 3, 2022, from https://www.slc.gov/parks/urban-forestry/urban-forestry-utah-native-trees/

3. *Logan, UT*. (n.d.). Loganutah.Org. Retrieved April 3, 2022, from https://www.loganutah.org/government/departments/environmental/forestry/index.php

4. Hennon, P. E., Frankel, S. J., Woods, A. J., Worrall, J. J., Norlander, D., Zambino, P. J., Warwell, M. V., & Shaw, C. G., III. (2020). A framework to evaluate climate effects on forest tree diseases. *Forest Pathology*, *50*(6). https://doi.org/10.1111/efp.12649

5. Rumpf, T., Mahlein, A.-K., Steiner, U., Oerke, E.-C., Dehne, H.-W., & Plümer, L. (2010). Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture*, *74*(1), 91–99. https://doi.org/10.1016/j.compag.2010.06.009

6. Abdulridha, J., Batuman, O., & Ampatzidis, Y. (2019). UAV-based remote sensing technique to detect citrus canker disease utilizing hyperspectral imaging and machine learning. *Remote Sensing*, *11*(11), 1373. https://doi.org/10.3390/rs11111373

7. Pérez-Bueno, M. L., Pineda, M., Vida, C., Fernández-Ortuño, D., Torés, J. A., de Vicente, A., Cazorla, F. M., & Barón, M. (2019). Detection of white root rot in avocado trees by remote sensing. *Plant Disease*, *103*(6), 1119–1125. https://doi.org/10.1094/PDIS-10-18-1778-RE

8. Deng, X., Tong, Z., Lan, Y., & Huang, Z. (2020). Detection and location of dead trees with pine wilt disease based on deep learning and UAV remote sensing. *AgriEngineering*, *2*(2), 294–307. https://doi.org/10.3390/agriengineering2020019

9. Thomas, S., Kuska, M. T., Bohnenkamp, D., Brugger, A., Alisaac, E., Wahabzada, M., Behmann, J., & Mahlein, A.-K. (2018). Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective. *Journal of Plant Diseases and Protection: Scientific Journal of the German Phytomedical Society (DPG)*, *125*(1), 5–20. https://doi.org/10.1007/s41348-017-0124-6

10. NAIP imagery. (n.d.). Usda.Gov. Retrieved April 3, 2022, from https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/index

11. Using Color Infrared (CIR) Imagery. Files.Nc.Gov. Retrieved April 3, 2022, from https://files.nc.gov/ncdit/documents/files/Using-Color-Infrared-Imagery-20110810.pdf

12. *How to open and work with NAIP multispectral imagery in R*. (2017, February 22). Earth Data Science - Earth Lab. https://www.earthdatascience.org/courses/earth-analytics/multispectral-remote-sensing-data/naip-imagery-raster-stacks-in-r/

13. Pant, R. (2020, December 1). *Accuracy and its shortcomings: Precision, Recall to the rescue*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2020/12/accuracy-and-its-shortcomings-precision-recall-to-the-rescue/

14. Lowe, A., Harrison, N., & French, A. P. (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods*, *13*, 80. https://doi.org/10.1186/s13007-017-0233-z

15. Nguyen, C., Sagan, V., Maimaitiyiming, M., Maimaitijiang, M., Bhadra, S., & Kwasniewski, M. T. (2021). Early detection of plant viral disease using hyperspectral imaging and deep learning. *Sensors (Basel, Switzerland)*, *21*(3), 742. https://doi.org/10.3390/s21030742

16. Viinikka, A., Hurskainen, P., Keski-Saari, S., Kivinen, S., Tanhuanpää, T., Mäyrä, J., Poikolainen, L., Vihervaara, P., & Kumpula, T. (2020). Detecting European Aspen (Populus tremula L.) in boreal forests using airborne hyperspectral and airborne laser scanning data. *Remote Sensing*, *12*(16), 2610. https://doi.org/10.3390/rs12162610

17. *Train Support Vector Machine Classifier (Spatial Analyst)*. (n.d.). Arcgis.Com. Retrieved April 4, 2022, from https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-analyst/train-support-vector-machine-classifier.htm

18. *Train K-Nearest Neighbor classifier (spatial analyst)*. (n.d.). Arcgis.Com. Retrieved April 4, 2022, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/train-k-nearest-neighbor-classifier.htm

19. Weaver, J., Moore, B., Reith, A., McKee, J., & Lunga, D. (2018). A comparison of machine learning techniques to extract human settlements from high resolution imagery. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*.

20. Taubenbock, H., Esch, T., Felbier, A., Roth, A., & Dech, S. (2011). Pattern-based accuracy assessment of an urban footprint classification using TerraSAR-X data. *IEEE Geoscience and Remote Sensing Letters : A Publication of the IEEE Geoscience and Remote Sensing Society*, *8*(2), 278–282. https://doi.org/10.1109/lgrs.2010.2069083

21. Calculate Confusion Matrices *(Calculate confusion matrices)*. (n.d.). l3harrisgeospatial.com. Retrieved April 24, 2022, from https://www.l3harrisgeospatial.com/docs/CalculatingConfusionMatrices.html.