INVASIVE SPECIES DETECTION FROM RGB AERIAL IMAGERY: INVESTIGATING LINKS BETWEEN PLANT CHARACTERISTICS AND TRANSFER LEARNING SUCCESS

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAIʻI AT HILO IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

TROPICAL CONSERVATION BIOLOGY AND ENVIRONMENTAL SCIENCE

DECEMBER 2021

By

Erica Ta

Thesis Committee:

Ryan Perroy, Chairperson

Travis Mandel

David Benitez

Jonathan Price

Keywords: invasive species, object detection, deep learning, transfer learning, Hawaiʻi

Acknowledgements

There are many people I have had the pleasure of collaborating with and learning from throughout this project. For that, I sincerely thank all those involved along the way. Firstly, I'd like to recognize my family's nonstop support and encouragement in all my pursuits thus far, especially when most opportunities have been far away from home. A big mahalo nui loa goes to Ryan Perroy for his contagious excitement and encouragement of geospatial science that started it all from the very first introductory call. I am truly grateful for his guidance in addition to all the connections and learning experiences he has facilitated since then. Another tremendous mahalo goes to Travis Mandel who has generously shared so much expertise on the topic of artificial intelligence with the UHH community and across various disciplines. It has been extremely rewarding to have learned about data science and machine learning through his passion for it. Mahalo to David Benitez for his dedication as a resource manager and instilling the significance of managing invasive species through my research project. Mahalo to Jonathan Price for providing his vast ecological knowledge towards this research that occupies a unique, interdisciplinary space. I am honored to have collaborated with so many experts in their respective fields. I collectively thank all my committee members for offering their help, time, effort, and great expertise to ensure the success and quality of this project as well as my manuscript. I cannot address the success of this project without acknowledging and deeply thanking the steadfast work of Pat Pérez. She was the backbone of this production, providing her intelligence and ever-present emotional support that has been invaluable for me and the Spatial Data Analysis and Visualization (SDAV) lab. Additionally, I'd like to thank Timo Sullivan and Danny Duda for their generous help during the data collection phase of this project. Another thank you to Max Panoff for jumpstarting all the deep learning pipelines I relied on for this project. Finally, to my family of friends and support systems I have been so lucky to be part of, mahalo nui loa for helping me in all kinds of ways to complete this endeavor.

The technical support and advanced computing resources from University of Hawai‘i Information Technology Services – Cyberinfrastructure, funded in part by the National Science Foundation MRI award #1920304, are gratefully acknowledged.

Abstract

Advancements in remote sensing techniques and deep learning applications like object detection have improved invasive species monitoring systems. Deep learning typically uses large training sets, up to millions of images, to consistently recognize targets, but generating these training sets may not be practical for incipient invasive species targets of interest. When large datasets are unavailable for training, one approach is to use transfer learning to overcome data limitations. The process applies knowledge learned from the source network (pre-trained on a task on which large datasets are available) to a target problem that has limited data samples. Here I examine how object detection performance for the following invasive species of interest in Hawai‘i differs with the inclusion of cross-species transfer learning: Miconia (*Miconia calvascens*), Guinea grass (*Megathrysus maximus*), and four symptomatic visible classes of Rapid ‘Ōhi‘a death (ROD)*:* red, brown, fine white, and skeleton. I also measured visual plant features of contrast, shape, size, and texture to understand how different plant morphologies provide easier or more challenging scenarios for plant object detection using aerial visible imagery. I found that 9 out of 30 transfer learning instances had significantly higher mean average precision (mAP) scores than instances without transfer learning ($p < 0.00167$, $\alpha = 0.00167$ (0.05/30). Transfer learning was found to be most effective between red, brown, fine white, and skeleton ROD classes and least effective among miconia and guinea grass. The feature measurement of contrast was significantly correlated with source model mAP ($R = 0.82$, $p = 0.045$) whereas texture was strongly correlated ($R = 0.77$, $p = 0.073$), size ($R = 0.54$, $p = 0.27$) was moderately correlated, and circularity ($R = -0.096$, $p = 0.86$) was weakly correlated. My results indicate there are advantages in plant detections that utilize similar source and target candidates for transfer learning, in addition to incorporating source candidates whose image data exhibit higher contrast and higher texture measurements. Overall, this study may inform future workflows to detect plants from aerial imagery by demonstrating how available data can be best leveraged or repurposed through transfer learning to detect a plant target using limited available datasets.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. Introduction

Hawaiian plant communities are notoriously susceptible to invasion by nonnative plants (Adam et al. 2014) and foreign pathogens (Mortenson et al. 2016). As Hawai‘i’s limited natural resources are ecologically impacted by ecosystem-modifying invasive plants, frequent monitoring and detection is critical for invasive species management. The prolific abundance of ecosystem-modifying invasive species in the Hawaiian islands has provoked efforts to increase systematic monitoring, detection, and eradication to manage invasive species populations (Cordell 2021; Fung Associates, Inc. & SWCA Environmental Consultants 2019).

Advancements in novel biodiversity monitoring systems are being developed from an array of *in situ* and remote sensing data to systematically assess change for multiple taxa (Anderson 2018) and can be applied to address invasive plant community distributions that occur over large extents. Advancements in remote sensing techniques have facilitated improved invasive species monitoring systems and broadened ecosystem analyses to larger scales, resulting in continuously rising quantities of high-resolution remotely sensed imagery. Invasive plant species identification from aerial imagery has previously been based on the use of human observation to directly count or digitize target species within an area of interest. This method is straightforward but has the disadvantage of being labor-intensive because all targets must be manually counted or digitized. Although remotely sensed monitoring has transformed the speed, scale, and cost of collecting ecological data, data analysis at the landscape scale still often involves costly, laborious human efforts to annotate or identify targets in large quantities of data. To optimize these drawbacks, deep learning techniques are being applied to image identification tasks across ecology (Borowiec et al. 2021; Brodrick et al. 2019; Rodriguez et al. 2021).

In recent years, applications of deep learning are becoming ubiquitous in automated image processing and are gaining application in large-scale environmental analysis with various tasks such as image classification, object detection, and image segmentation (Brodrick et al. 2019; Weinstein et al. 2019; Oghaz et al. 2019; Guo et al 2016). Deep learning has also been widely adopted for detecting and identifying objects from RGB imagery and is being increasingly applied to plant identification (Rodriguez et al. 2021; Kattenborn et al. 2020; Weinstein et al. 2019; Csillik et al. 2018). Convolutional neural networks (CNN), one of the

most extensively used deep learning models, commonly use large training sets of thousands of segmented images to learn to recognize targets because there is considerable variability in real-world settings (Krizhevsky et al. 2017). This study focused on object detection, determining where objects are located in an image (object localization) and which category each object belongs to (object classification) (Zhao et al. 2019). Object detection involves two major components of training models and evaluating models. Training an object detection model involves providing example annotations from images on which the network learns. Second, models are then evaluated on how well the model can generalize or detect the object of interest in photos it has not seen. The inputs used to train an object detection network are rectangular annotations that outline the target of interest in a photo also known as bounding boxes. The bounding boxes are input as examples and are used in training the models as well as model evaluation. For evaluating models, ground truth bounding box annotations are compared against the predicted bounding boxes.

One of the approaches to convolutional neural network-based object detection involves a one-stage detector (Sultana et al. 2019). Examples of one-stage detectors include YOLO (Redmon et al. 2016), SSD (Liu et al. 2017), and RetinaNet (Lin et al. 2018), that use anchors or sets of pre-defined bounding boxes of varying scales and ratios for the initial region proposals and the detector classifies these pre-defined regions (Rodriguez III et al. 2021). This study selected a network based on a one-stage RetinaNet model (Lin et al. 2018) because of its improved performance over other detectors for tree crown detection (Culman et al. 2020; dos Santos et al. 2019). While object detection shows promise in tree crown detection and transfer learning from publicly available datasets, little is known on whether transfer learning between source and targets of plants in aerial imagery helps significantly improve detection performance.

To help relieve data limitation challenges, transfer learning has been applied in deep learning systems (Yosinski et al. 2014). In transfer learning, a source network is first trained on a relatively large dataset, then those weights are used as a starting point when training another task with a smaller dataset (Yosinski et al. 2014). A common method of transfer learning uses a single level of transfer learning where pre-trained models from large, publicly available image datasets for the source task such as *MSCOCO* or *ImageNet*, which contain tens of thousands to millions of labeled images of many categories (Krizhevsky et al. 2012; Kattenborn et al. 2020).

However, these commonly used datasets contain everyday objects like bicycles, pets, and cars usually taken from a personal camera or smartphone, while the data collected to monitor and detect invasive plants in Hawai‘i are aerial imagery captured from a bird’s eye view. The content within these aerial images often looks very different from images and categories available in existing large datasets, such as ImageNet. Because commonly used datasets like ImageNet do not contain labeled aerial imagery examples, additional data sources are needed to better suit detection of plants found in aerial imagery.

Hawai‘i’s high degree of vulnerability stemming from widespread populations of invasive species present many opportunities where deep learning methods should be applied to detect invasive plants more efficiently. However, there will be an ongoing challenge to acquire sufficient imagery and labeled image datasets when detecting incipient populations of invasive plants in Hawai‘i. Additionally, because some plants of interest are endemic to Hawai‘i, searching the internet to find additional labeled imagery is impractical. Despite their great potential, real-world applications of deep learning methods are commonly constrained by small training datasets that ultimately restricts final performance (Ma et al. 2019).

Aerial imagery collected during invasive plant or pathogenic species monitoring surveys is a growing but still relatively scarce source of species-specific data in Hawai‘i, especially for incipient populations of newly detected target species of concern. As transfer learning has demonstrated a productive approach to mitigate data deficiencies, mostly in object recognition tasks, the most appropriate method of transfer learning will vary by objective since there is no single guide to transfer learning (Day & Khoshgoftaar 2017). I was interested to explore if a second transfer learning step using aerial plant imagery can be used to improve our ability to identify non-native plant species from aerial imagery in Hawai‘i with object detection. Rather than spending additional resources to collect more imagery on a particular species lacking data, one objective of this study is to apply transfer learning as a method of determining how to best bootstrap deep learning performance. The goal is to develop a system to automatically inform researchers which existing plant dataset is most useful to incorporate in transfer learning to improve object detection performance of a new and different plant target of interest.

To better understand the kind of data that facilitates improvements in object detection and transfer learning, it is important to consider the various features within an image. Feature

extraction constitutes one of the pillars of the identification and the classification of objects based on computer vision (Perez-Sanz et al. 2017). Deep learning algorithms have successfully managed to extract feature information automatically, but it is challenging to tell directly what features the neural networks are using for its computation to make predictions (Sestili 2018). Due to the feature richness of high-dimensional convolutional filters and the large number of connections between hidden layers in neural networks, the interpretation of these learned features is difficult (Horwath et al. 2020). Instead, considering image features commonly used in remote sensing may help determine the kinds of data that would result in detection improvements via transfer learning. Some common features used to identify objects in an image are edges, intensity of image pixels, geometries, textures, and colors (Perez-Sanz et al. 2017). A previous study demonstrated that shape alone was insufficient for plant identification and other aspects, such as color and texture, should be incorporated to increase performance of identification system (Kadir et al. 2011). Given this information, a suite of visual features was selected to measure and then evaluate the performance of species identification based on these plant features.

This study proposed to measure how object detection performance for three targets of interest in Hawai‘i (Miconia (*Miconia calvascens*), Guinea grass (*Megathrysus maximus*), and Rapid ‘Ōhi‘a death (ROD) (comprised of two fungal pathogens, *Ceratocystis lukuohia* and *Ceratocystis huliohia)* (Barnes et al. 2018) differs when each source model is used to pretrain a different target’s model with limited training examples. Comparing the output detection model performance per species, with and without transfer learning, and comparing these results with measured plant morphometric parameters, may indicate which plant morphologies and features provide easier or more challenging for object detection of plants in aerial visible imagery. I compared CNN model performance for the detection of guinea grass, miconia and four ROD symptomatic classes: red, brown, fine white, and skeleton, with and without transfer learning. Another study objective was to understand whether plant features in aerial high-resolution visible imagery affect detection performance and whether these features can predict the efficacy of transfer learning. Visual plant features of color, shape, size, and texture were measured to determine if there is a connection between detection performance and the features (Kadir et al. 2011).

I hypothesized that transfer learning would produce the biggest improvement in detection accuracy when Guinea grass detection models are used to train other target models, as guinea grass exhibits a visual combination of shape irregularity and high contrast when it appears on darker backgrounds, such as the lava fields in the guinea grass imagery used in my thesis. I anticipated these properties to represent distinctive features, therefore facilitating an easier ability to learn and produce the highest performing models (Kadir et al. 2011; Perez-Sanz et al. 2017). Based on previous success of models detecting coconut and pandanus trees, guinea grass as seen in an aerial image resembles the aerial view of coconut or pandanus most out of the suite of objects of interest in this study (Perroy et al. unpublished data). I suspected the guinea grass model may result similarly well in terms of detection performance. Because data used for guinea grass in this study exhibit contrasts from having a darker background and lighter foreground in addition to shape irregularities, I expected species with multiple distinct features would facilitate greater transferred knowledge to a different species.

A primary goal of this research was to determine which source model of six targets helps improve another target's detection model the most. Understanding how features from a particular target affect performance of another task's (or species') model may help inform future data collection efforts to monitor other invasive species in Hawai‘i. The objective was to identify whether cross-species transfer learning for all six target datasets with one another can develop a better detection system than if one target was trained on its own. Moreover, understanding model performance based on plant features may help us better understand why transfer learning object detection performs better or worse with different target plant species. The results of this approach may illustrate applications where deep learning is most effective in Hawaiian ecological conservation as well as identify different, but related, types of data that might be most helpful. Future ecological conservation work may benefit from what is learned about automated detection methods by diminishing costs of laborious human identification and reducing efforts involved in collecting new data through leveraging existing plant data based on shared features.

II. Methods

A. Overview

The general methodology involved two main components: measurements to characterize a suite of four plant features for each target species and then object detection with transfer learning for the six targets (Figure 1). The process of feature measurements and object detection was replicated for each target of interest. In this study, models were first pre-trained using the ImageNet dataset, but underwent a second transfer learning process by using pre-trained networks from individualized plant models trained with significantly less data than available in the ImageNet dataset. To elaborate, this study utilized Google TensorFlow as the machine learning platform, Keras API, RetinaNet as the object detection model, and ResNet-50 for the backbone (Figure 2).

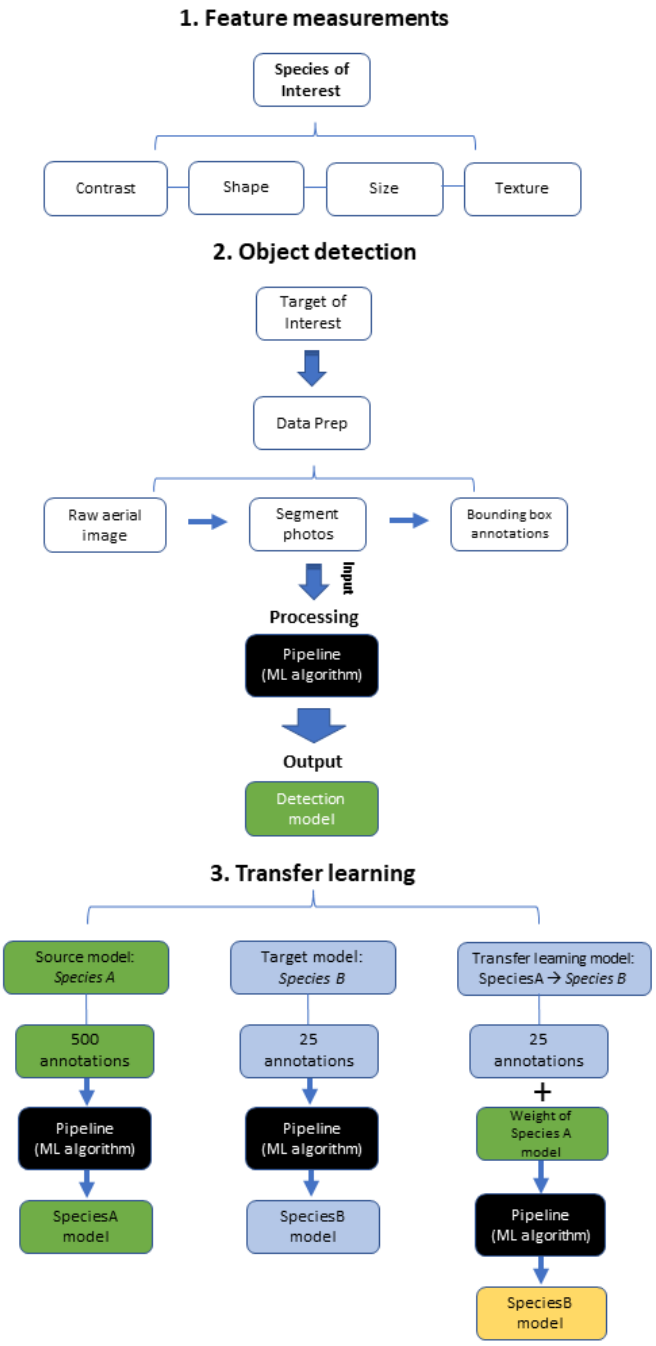

Fig. 1. Workflow illustration of methodology categorized into two parts. This methodology was replicated for each species of interest. The first task shows feature measurements of Contrast, Shape, Size, and Texture. The second part describes object detection, beginning with image pre-processing where all required photos are segmented into 1,000 x 1,000 pixel images. Photo datasets are manually annotated for species of interest, then processed through the object detection algorithm to produce a detection model. Transfer learning is then performed and utilizes weight from a detection model from one target (using 500 annotations) and reutilizes it to train a different target (using 25 annotations and a pretrained model).

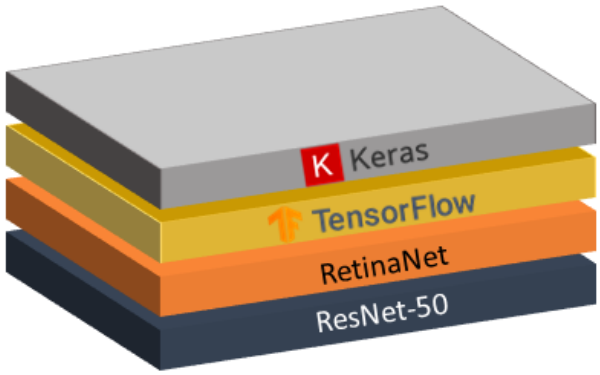

Fig. 2. Visualization of this study's object detection framework: Keras as the Application Programming Interface (API), Google TensorFlow as the machine learning platform, RetinaNet as the object detection model, and ResNet-50 for the backbone.

All datasets were preprocessed by segmenting full size images into 1,000 x 1,000 pixel size images and bounding box annotations of target species as input. There were varying number of annotations per each target of interest, but at least 715 annotations. The 715 annotations per target of interest were split across three datasets: 70% of annotations went to training, 20% went to validation, and 10% in the test dataset. Each dataset included images and bounding box annotations that were created using LabelImg. Each target species model was trained and resulted in individual detection models also known as inference models. Three series of training occurred to compare models used with and without transfer learning: source models trained with 500 annotations, target models trained with 25 annotations, and transfer models trained with 25 annotations plus another target's pretrained model (Figure 1). 500 annotations (70% of 715) were reserved for training source models, then 25 annotations were chosen for target and transfer learning models to simulate limited data scenarios. The amount of 25 annotations were chosen because miconia was found to detect well using around 30 annotations in experimental detection models for miconia. For transfer learning, each target's detection model was applied to pretrain a different target equipped with less training data. Any given model's performance was measured using Precision and Recall (PR), where precision is the number of true positives over the number of predicted positives and recall is the number of true positives over the number of actual positives. Precision and recall are provided at a specified confidence threshold. In addition, the mean average precision (mAP) is a single evaluation metric that integrates over all confidence

thresholds and was selected as the model performance metric in this study. The intersection over union (IoU) is the intersection between the predicted bounding box and actual ground truth bounding box. As the IoU decides whether a prediction is correct with respect to an object or not, the mAP was derived from an IoU of 0.5 or 50% overlap between ground truth and predicted boxes. The ground truth boxes and predicted boxes were matched by taking coordinates from the ground truth annotations and the coordinates of the predicted annotations to calculate the IoU score. As multiple predicted boxes corresponded to the same ground truth box, the predicted box with the highest confidence prediction score would be matched with the ground truth box. Precision and Recall are provided at a certain confidence threshold whereas mean average precision (mAP) integrates over all confidence thresholds and is represented as a single evaluation metric (Padilla et al. 2020). The final transfer learning analysis compared performance between two series: the source series trained with 500 annotations and the transfer learning series trained with only 25 annotations but included the addition of another target's pretrained model. The performance of all models (mAP) and feature measurements (contrast, shape, size, and texture) were analyzed to reveal any correlation between performance and plant traits. The following sections elaborate on the elements and steps of this study's methodology.

B. Targets of interest

Three ecosystem-modifying invasive species were included in this study: Guinea grass (*Megathrysus maximus*), Miconia (*Miconia calvascens*), and Rapid ʻŌhiʻa Death (ROD) (comprised of two fungal pathogens, *Ceratocystis lukuohia* and *Ceratocystis huliohia*), categorized in this study as four distinct classes representing its symptomatic stages: red, brown, fine white, and skeleton (Figure 3).

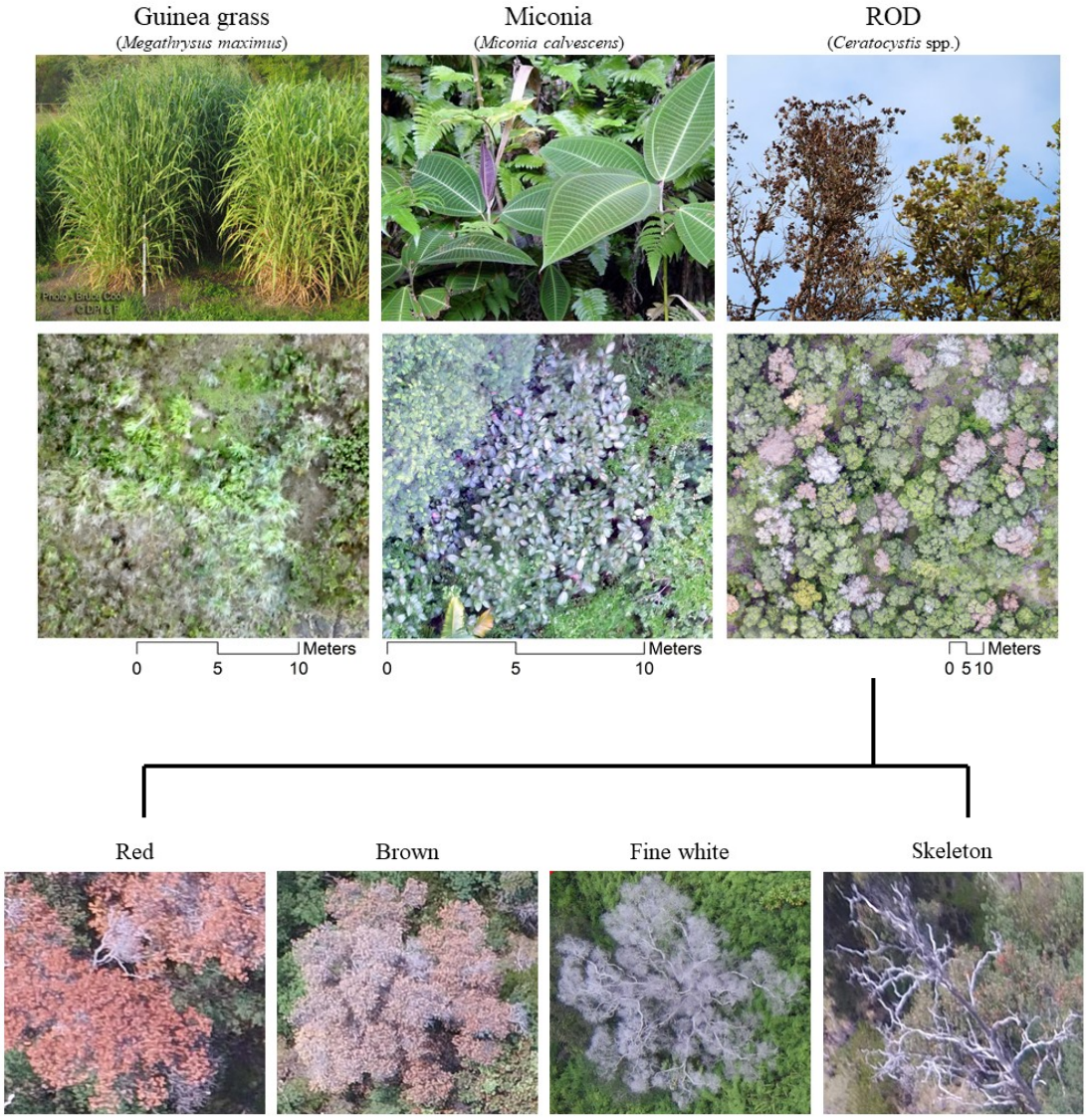

Fig. 3 Each species of interest seen on ground (first row) and in aerial imagery (second row). Below are classes of red, brown, fine white, and skeleton illustrating symptomatic stages of ROD progression from initial leaf discoloration to loss of leaves. Guinea grass ground photo provided by Bruce Cook, miconia ground photo provided by Big Island Invasive Species Committee, and

ROD ground photo provided by Hawai‘i Volcanoes National Park. All other imagery provided by Spatial Data and Visualization Lab at University of Hawai‘i at Hilo.

1. Guinea grass (*Megathrysus maximus*)

Guinea grass is among one of the most pertinent invasive, exotic grasses in Hawai‘i, Brazil, Australia, and Taiwan, among others (Ellsworth et al. 2013; Mantoani et al. 2016; Klinken & Friedal 2017; Ho et al. 2015). It has adapted to establish in ecosystems ranging from dry to wet conditions, and in turn, leads to increased flammability by increasing the fuel load and continuity within the invaded habitat (Ellsworth et al. 2013). This grass has high environmental adaptability known to slow ecological succession of plants located in its spatial distribution. Guinea grass is considered a top invasive species of concern in Hawai‘i Volcanoes National Park (HAVO) (Benitez et al. 2012). This native African species is a robust C4 bunchgrass that is commonly found in the park's coastal lowlands and sub-montane seasonal zones (Benitez et al. 2012).

2. Miconia (*Miconia calvascens*)

Miconia is native to rain forests from southern Mexico to southern Brazil and northern Argentina where it occupies a wide altitudinal range from lowland to mountain tropical forests (up to 1,800 m) (Gonzalez-Munoz et al. 2015). However, miconia was designated by the IUCN Invasive Species Specialist Group as one of the "100 worlds invasive alien species" (Lowe et al. 2000) and has highly invaded Hawai‘i, French Polynesia, the region of Queensland in Australia, Sri Lanka, and New Caledonia (Gonzalez-Munoz et al. 2015). This species causes harmful impacts on native flora because its large, broad leaves decrease the amount of light that reaches the soil as it forms dense monospecific stands that suppress native vegetation (Leary et al. 2018; Meyer et al. 2011).

3. Rapid ʻŌhiʻa Death (ROD)

ʻŌhiʻa lehua (*Metrosideros polymorpha*) is Hawaiʻi's most widespread and culturally significant native tree. This foundational species is essential for native Hawaiian forest ecosystem function, provides critical habitat for other native plant and animal species, and majority of Hawaiʻi's freshwater resources ((Loope et al. 2016; Asner et al. 2018). In 2009-2010, an alarmingly widespread mortality of ʻōhiʻa emerged on Hawaiʻi Island. The manner of tree death was unique as mortality began with a branch, fork, or an individual tree completely dying before adjacent trees showed signs of stress ((Mortenson et al. 2016). Commonly known as Rapid ʻŌhiʻa Death (ROD), this process is distinct from from "ʻōhiʻa dieback", episodes of naturally occurring cohort dieback that happens at a landscape scale (Loope et al. 2016). ROD is characterized by two fungal pathogens, *Ceratocystis lukuohia* and *Ceratocystis huliohia* (Perroy et al. 2021). Despite pathogen identification, the cause of ROD has not been determined, but has been hypothesized to be due to a combination of anthropogenic movement of infected material used for firewood, contaminated tools, aerial dispersal of infected insect boring dust, and recently linked to ungulates (Barnes et al. 2018; Perroy et al. 2021). Rapid ʻōhiʻa death is an ongoing threat to ʻōhiʻa forests statewide.

Both fungal species cause browning foliage that is then followed by rapid tree mortality due to fungal growth blocking transport of water and sugar throughout the vascular system (Barnes et al. 2018; Fortini et al. 2019). There is a consistent progression and characteristics of ROD symptomatic trees as the fungal infection within an ʻōhiʻa tree moves from initial to final stages of the disease's expression. Therefore, suspect ROD progression seen in aerial imagery was categorized into four main classes: (1) Red, (2) Brown, (3) Fine white, and (4) Skeleton (Figure 3) (Perroy et al. 2021).

C. Plant feature indices

To understand whether plant features affect object detection performance and if these features can predict the efficacy of transfer learning, Table 1 lists plant features commonly used in automated image processing analyses like image recognition and image segmentation (Kadir et al. 2011; Perez-Sanz et al. 2017). Other plant identification features were considered for this study, but the selected features below reflect plant features most suspected to have an influence

in an object detection task because they were found to be highly influential in automated plant identification tasks at the leaflet scale (Jamil et al. 2015; Babatunde et al. 2015). These characteristics are also commonly used for manual interpretation of ecological features in an aerial photography (Morgan et al. 2010). Plant characteristics may address whether object detection performance of plant species and pathogens can be explained by different traits from this study's species of interest. Each feature category was measured with a sample of 50 individuals per target of interest (Figure 4). The 50 individuals were selected by first taking the bounding box areas from the "Train" datasets of each target of interest, cropping out each bounding box area, and saving it as an individual image. 50 images were then randomly selected per target of interest. Following, each target of interest was outlined using ImageJ, a public domain Java image processing program, and was used to calculate all feature indices (Schneider et al. 2012). Each target of interest was outlined within its bounding box to better capture and isolate each target's unique feature characteristics rather than taking feature measurements on the whole bounding box (Figure 4).

Table 1. Feature indices to be quantified per each species of interest.

Characteristic	Index	Procedure	Source
Contrast	Root-mean-square (RMS) contrast	$\sqrt{\frac{1}{N}\sum_{j=1}^{N}\frac{(L_j-\bar{L})^2}{\bar{L}}}$	Triantaphillidou et al. 2019
Shape	Circularity	4π *Area / Perimeter^2	Russ & Russ 2017; Olson 2011
Size	Pixel area	Area in square pixels from 1000 x 1000 pixel image	Schneider et al. 2012
Texture	Variance	$\sum_{i=0}^{N_g-1}(i-M)^2\ P(i)$	Exelis Visual Information Solutions 2010

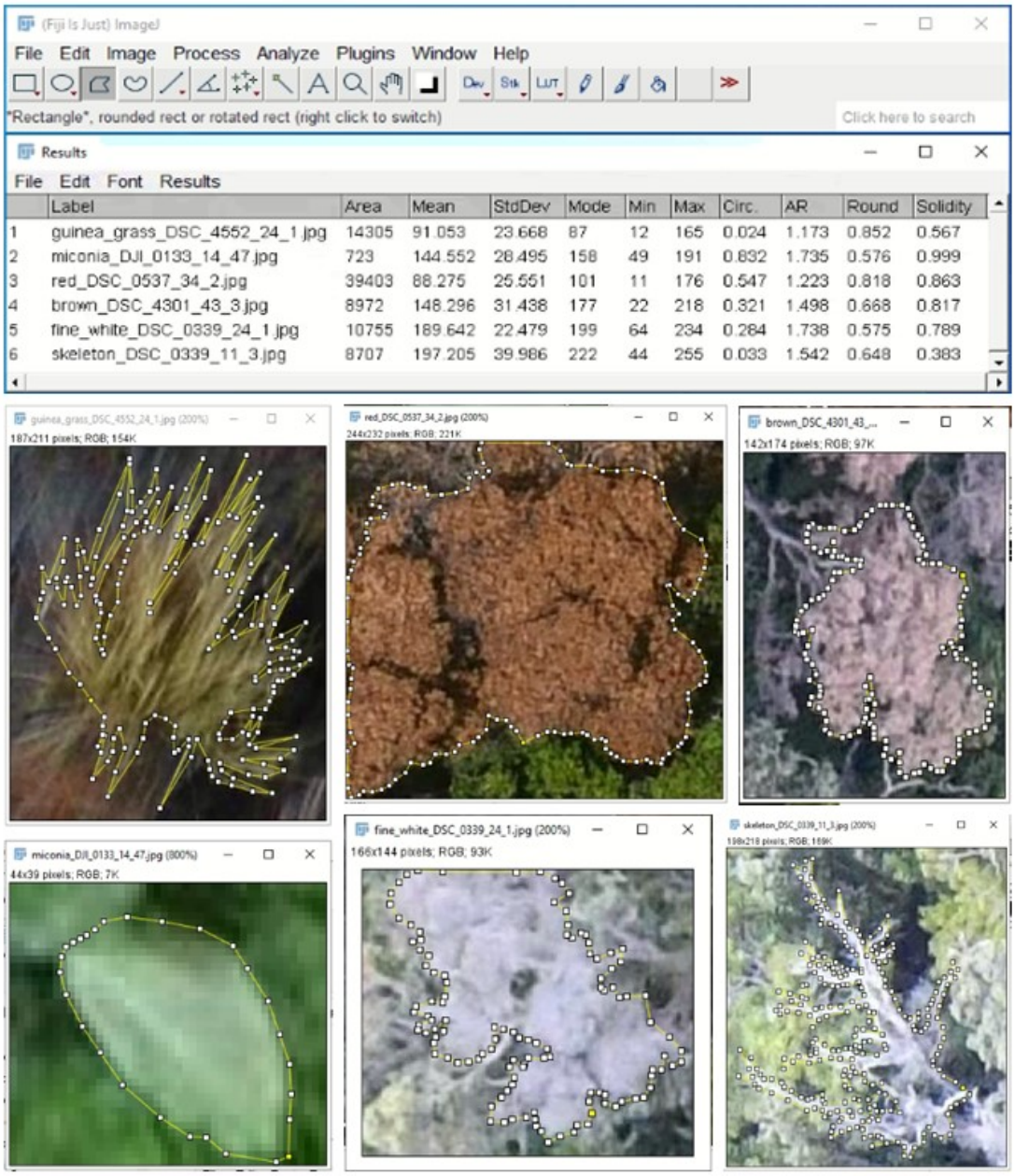

Fig. 4. Illustration of feature measurements collected in the ImageJ software such as area, standard deviation, and circularity (Circ.).

a. Contrast

The basic perceptual attribute of an image is contrast, also a measurement of the human visual system sensitivity (Bhuiyan & Khan 2018). Perceptual contrast can be defined as a perceived luminance variation (Triantaphillidou et al.

2019). Natural images contain substantial spatial variation in both local contrast and local luminance (Frazor & Geisler 2006). This study used root-mean-square (RMS) contrast as the image contrast measurement that is equivalent to the standard deviation of luminance (Moulden et al. 1990; Frazor & Geisler 2006; Bhuiyan & Khan 2018; Triantaphillidou et al. 2019). Below shows the C_{RMS} expression below where N is the number of pixels in the whole image, L_j is the displayed luminance of the jth pixel and $\bar{L}$ is the mean luminance:

$$C_{RMS} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\frac{\left(L_j-\bar{L}\right)^2}{\bar{L}^2}} \quad (1)$$

b. Shape

As shape refers to general form or outline of individual objects (Blaschke et al. 2014), plants exhibit a variety of branching patterns and foliage arrangements where these shape features have been leveraged to improve or guide automated image analysis tasks, such as object segmentation. Shape of plants can be used as learned properties of an object to increase needed information that facilitates segmentation (Borenstein & Malik 2006). This study used Circularity, a measure of roundness (area-to-perimeter ratio), obtained as the ratio of the area of an object to the area of a circle with the same convex perimeter (Russ & Russ 2017; Olson 2011). All circularity measurements were taken from aerial imagery.

While this measure of circularity provides some sense or shape in terms of its difference from roundness, certain plant outlines such as grasses or a tree's skeleton are not round but can exhibit an overall round pattern. As a result, this metric does not thoroughly distinguish between the distinctive shapes of plants which is a limitation of the circularity metric used in the context of this study.

$$Circularity = \frac{4\pi \cdot area}{(convex\ perimeter)^2} \quad (2)$$

c. Size

The size of an object in an image can be defined as the number of pixels the object occupies. Size is utilized in manual interpretations of aerial imagery to determine ecological features such as vegetation age, structure, and species identification. The relative and absolute size of objects in an image is used to make ecological inferences and identify natural features (Morgan et al. 2010). Size measurements were included in this study to determine if scale is accounted for in object detection. Various plant species in Hawaiian forests exhibit similar leaf shapes, but of different sizes. Quantifying size of each species of interest will help illustrate if object detection is utilizing differences in size during its detection process. For this study, the size of an object was constituted by the area of pixels an object represents in a 1000 x 1000 pixel image.

d. Texture

Image texture is referred to as frequencies of change in tones and their resulting spatial arrangements (Blaschke et al. 2014). The essence of texture in digital imagery is a visual pattern consisting of entities with certain color, shape, size, etc. characteristics. These characteristics can give the perceived coarseness, smoothness, randomness, uniformity, which can be regarded as texture (Yuan et al. 2019). Texture analysis is key in common applications of image segmentation, image classification, and pattern recognition. Texture was measured by statistical variance of pixel values of a region of interest per each species of interest. Below is an expression of variance as a measure of the dispersion of values around the mean where *P(i)* is the probability of each pixel value and N_g, the number of distinct grey levels in the quantized image:

$$Variance = \sum_{i=0}^{N_g-1} (i - M)^2 \; P(i) \tag{3}$$

D. Deep learning

Deep learning is introducing new methods for remote sensing analysis in regard to image analysis and computer vision (Kattenborn et al. 2021). Deep learning is a type of representation-learning, a set of methods that allow a machine to be fed raw data and automatically discover the representations needed for detection or classification (Ball et al. 2017). Deep learning is characterized by a significantly increased number of successively connected neural layers, where each layer is known as a feature map. Neural networks with these deep structures refer to the high number of processing layers in which input data passes (Zhao et al. 2019). As a result, complex features can be learned. Deep learning algorithms are capable of automatically extracting feature information, but determining which features the neural networks used for its computation to make predictions is not very interpretable (Sestili 2018). Although, in certain applications, researchers have observed that in the first layer of representation, the presence or absence of edges at certain orientations and locations in the image are determined. During the second layer, certain patterns are detected through perceiving particular arrangements of edges. In the third layer, patterns may assemble into larger combinations that correspond to parts of familiar objects. Then, subsequent layers would detect objects as a full combination of these parts (LeCun 2015).

E. Object detection with deep learning

a. Object detection

The objective of object detection is to develop computational models and techniques that answer: *What objects are where?* (Zou et al. 2019). Object detection was chosen for this study because it can identify which plant is in an image as well as extract geographic information on where the plant is located. The inputs for an object detection network are rectangular annotations that outline the target of interest in a photo also known as bounding boxes. The bounding boxes are input as examples and are used in

training the models as well as model evaluation. For evaluating models, ground truth bounding box annotations are compared against the predicted bounding boxes.

In this study, the object detection-based CNN was trained using bounding boxes of a single class of interest. Models are trained with a set of bounding boxes and images and are evaluated on a separate set of images and annotations the network has not seen. Once models are trained, results from the evaluation include a confidence (typically shown as a percentage) of how well the species of interest was correctly identified for each predicted bounding box (Brodrick et al. 2019).

b. RetinaNet for object detection

Recently improved object detection methods are built on trainable architectures that contain an increased number of successively connected neural layers to learn more complex features (Kattenborn et al. 2021). Recently, one-stage object detectors in deep learning have achieved matching performance with two-stage detector accuracy and speed with architectures such as RetinaNet, a one-stage detector composed of a backbone network and two task-specific subnetworks (Lin et al. 2018; Wang et al. 2019). RetinaNet was chosen for its improved performance for tree crown detection in comparison to other detectors (Culman et al. 2020; dos Santos et al. 2019). The RetinaNet architecture is named for its dense sampling of object locations in an input image (Lin et al. 2018). There are two main components in RetinaNet which is the backbone feature pyramid network (FPN) and focal loss. The backbone network is an off-the-shelf CNN which calculates feature maps at different scales and includes variants like ResNet-50, Resnet-101, or ResNet-152 (Ahmad et al. 2020; Lin et al. 2018). The backbone FPN is used to address various scales of target features for target species classification and location. After feature maps are extracted using the backbone network, convolutional object classification is performed on these features by using the first subnet. Bounding boxes are determined through a regression problem using the second subnetwork (Ahmad et al. 2020). Focal loss is used to address class imbalance and enforce importance of hard examples (Wang et al. 2019), which focuses training on a sparse set of hard examples and

prevents the vast number of easy negatives from overwhelming the detector during training (Lin et al. 2018). For this study, ResNet-50, a CNN that is 50 layers deep, was the primary backbone used for object detection.

F. Google TensorFlow

TensorFlow is an open-source software library for machine learning and artificial intelligence. It is an interface for expressing machine learning algorithms and is the underlying implementation for executing these algorithms (Abadi et al. 2016). TensorFlow controls low-level management of the learning algorithms.

i. Keras & Keras RetinaNet

Keras is a high-level neural networks Application Programming Interface (API), that is written in Python and capable of running on top of TensorFlow. Keras provides a high-level interface in Python. It is a deep learning library that supports convolutional networks and runs on CPU and GPU (Chollet 2015). Keras was used for its more user-friendly interface while using TensorFlow as a backend. More specifically, our pipeline used an existing open-source Keras implementation of RetinaNet (Gaiser et al. 2021).

G. Imagery

Two types of deployment methods were used to acquire imagery in this study via helicopter and drone. The helicopter was affixed with a camera rig on the belly of the helicopter with Nikon D850 cameras attached. The drone imagery was captured using a Zenmuse X5s camera. A drone was used to collect aerial imagery for miconia, helicopter was used for guinea grass imagery, and both drone and helicopter were used to collect ROD imagery. The elevation flown above ground level for each target of interest ranges from 30 – 300 meters. The resulting imagery resolutions across all targets of interest range between 1.5 – 5.5 cm. Specifications for all imagery collection are shown in Table 2 and Table 3.

Table 2. Camera specifications for all imagery used in this study.

Camera	Deployment Method	Sensor type	Resolution	Sensor size (mm)
DJI Zenmuse X5S	Drone	CMOS	5280 x 3956	20.89
Nikon D850	Helicopter	CMOS BSI	8256 x 5504	35.9 x 23.9

Table 3. Species of interest and their respective imagery specifications.

Species	Image location	Above Ground Level (m) (min - max)	Sampled resolution (cm)	Camera	Sensor size (mm)
Miconia	Various locations across Hawai‘i Island	30-50	1.5	DJI Zenmuse X5S	20.89
Guinea grass	HAVO	153	1.9	Nikon D850	35.9 x 23.9
ROD	Various locations across Hawai‘i Island	90-120	2.7	DJI Zenmuse X5S	20.89
ROD	HAVO	244-300	3.75-5.5	Nikon D850	35.9 x 23.9

H. Image preprocessing

a. Segmenting images

Full size imagery was segmented into 1000 x 1000 pixel segments (Figure 5). 1000 x 1000 square pixel segments were chosen to prevent a loss in fine-grained details. If a larger input is given, the RetinaNet pipeline will automatically resize the image to 1000 pixels per side, which would result in a loss of detail.

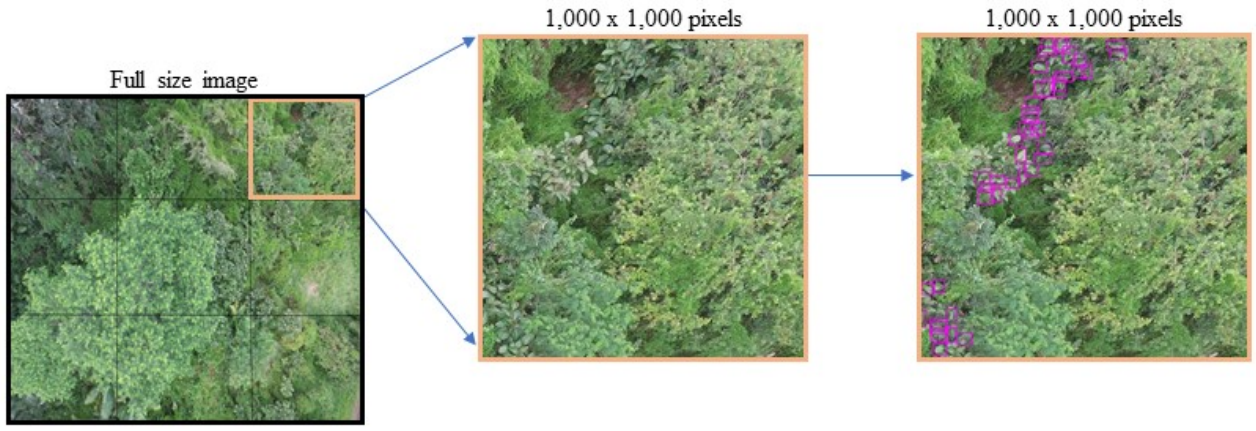

Fig. 5. Illustrates data processing stage where full size images are tiled into 1,000 x 1,000 pixel size images.

b. Bounding box annotations

LabelImg, a graphical annotation tool, was used to label object bounding boxes in images after they were segmented into 1,000 x 1,000 pixel size images (Tzutalin 2015). Each species was labeled as their common name, such as “miconia” or “guineagrass” (Figure 6).

Fig. 6. Example of creating bounding box annotations in LabelImg of classes red, brown, and fine white.

I. Training & evaluation

Once annotations were completed, a CNN was fed a selected target's training dataset. For this study's application, a dataset is composed of a set of corresponding images and annotations. At this stage, the network attempted to learn patterns in the provided training imagery and bounding box dataset. Every object detection method includes several hyperparameters or variables that are set before training occurs to determine the network structure and the variables that determine how the network is trained (Muthuraja et al. 2020). The hyperparameters that were applied during the training stage include data augmentation where photos were randomly transformed, batch size, step size, number of epochs, chosen backbone, and which layers were frozen (Table 4). The hyperparameters for all model series remained the same except the backbone was frozen to prevent overfitting on a smaller dataset for the target and transfer learning model series.

When training finished, a collection of snapshots or models were output that represent benchmarks throughout training. Snapshots were converted into inference models for evaluation and testing to identify targets from validation images that were held out of training. Because neural networks process data in a random order every time a model is run even with the same training data, each training instance was repeated five times with the same data to better generalize a model's performance of a given dataset. To compare model performance between each target with or without transfer learning, a series of source, target, and transfer learning models were created.

Table 4. Hyperparameters used for training per individual model series.

Model	Batch size	Step size	# of epochs	Backbone	Layers frozen
Source	3	167	100	ResNet-50	None
Target	3	167	100	ResNet-50	Freeze backbone
Transfer	3	167	100	ResNet-50	Freeze backbone

a. Transfer learning

In transfer learning, a source network, Target A, is trained with its own Target A dataset. Target A's model is then applied as a pre-trained model to then learn a second target network and task for Target B (Yosinki et al. 2014). Each target's individual source model was initially trained using ImageNet as base knowledge to generalize object detection in images. By using the "pre-trained" source models trained on 500 annotations, the weights were directly applied or transferred to train another target of interest with much less data at 25 annotations. In this study, the 500 annotations used in the source models can be referred to as the "source" annotations while the 25 annotations used in the target and transfer learning models can be referred to as the "target" annotations.

Images and annotations were taken from individual flights of varying amounts of images and annotations. To reach exactly 500 annotations per target of interest, some images and annotations were added or deleted from the training dataset. Models of the target series were trained using 25 target annotations, a fraction of the number of source model annotations. 25 annotations were chosen because experimental models ran to detect miconia resulted in high performance using 30 miconia annotations. Decreasing the annotations to 25 simulated a more limited data scenario for the target and transfer learning models. For transfer learning models, 25 annotations were used for training plus the addition of a pretrained model (Figure 7).

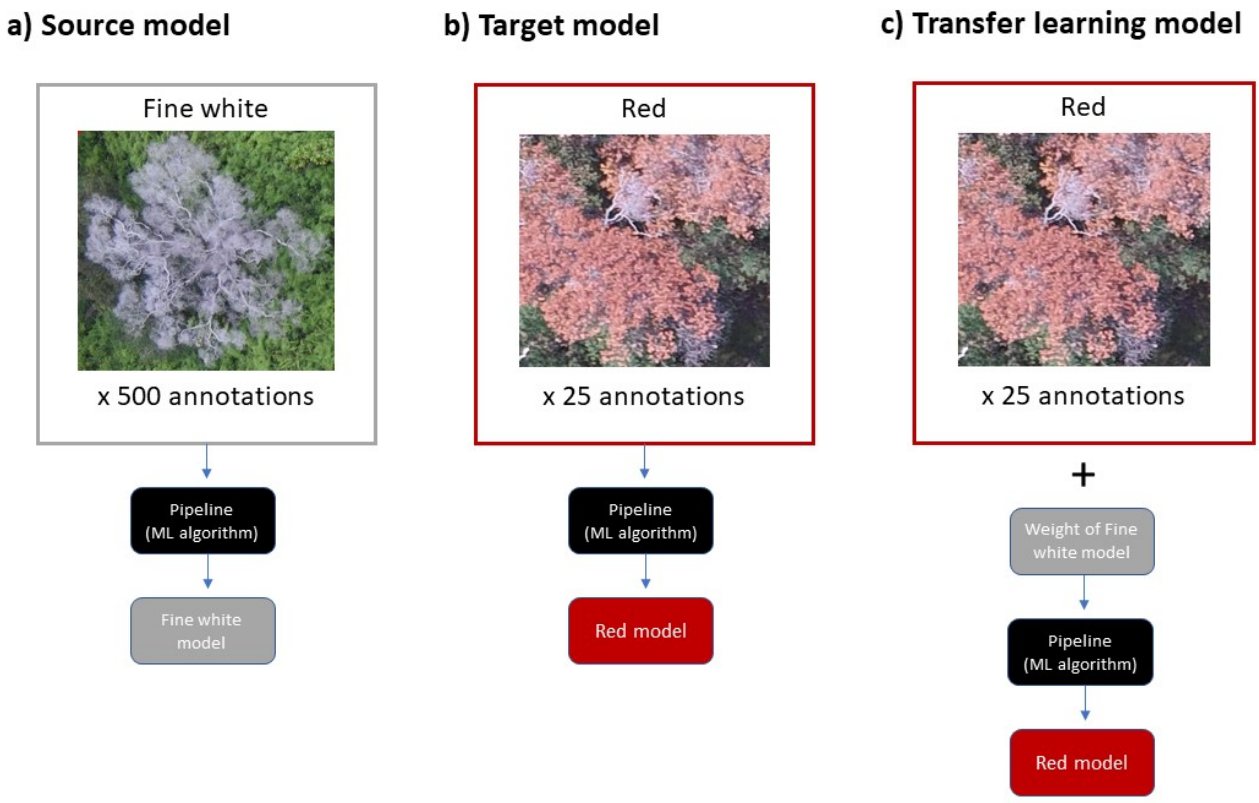

Fig. 7. Illustrates three series of models run with varied training conditions: source, target, and transfer learning series. A "fine white" source model was trained with 500 annotations, a "red" target model trained with 25 annotations, and a "fine white to red" transfer model trained with 25 annotations plus a pretrained "fine white" model from the source series.

b. Evaluation

Object detectors use multiple criteria to measure performance of a model. Object detection performance is typically measured from two metrics calculated in the confusion matrix in Table 5: precision $\left(\frac{TP}{TP+FP}\right)$ and recall $\left(\frac{TP}{TP+FN}\right)$. Precision measures the number of true positives over the number of predicted positives and recall is the number of true positives over the number of actual positives. Precision is derived from the Intersection over Union (IoU) which is the ratio of the area of overlap and the area of union between the ground truth and the predicted bounding box (Rezatofighi et al. 2019) (Figure 8). The ground truth boxes and predicted boxes were matched by taking coordinates from the ground truth annotations and the coordinates of the predicted annotations to calculate the IoU score. As multiple predicted boxes corresponded to the same ground truth box, the predicted box with the highest confidence prediction score would be matched with the ground truth box. This was done by sorting the confidence score from highest to lowest before any matching or comparison was made. Once a ground truth box was paired with a (higher confidence) predicted box, it is marked as a True Positive and that ground truth box is no longer in consideration for matches with lower confidence predicted boxes. It is important to note that the code used to match predictions to ground truth boxes greedily considers the matches in order after sorting by detection confidence. This is a limitation which may have a small impact on the resulting mAP scores.

For this study, an IoU threshold ratio of 0.5 was set to determine if a detection is correct. A prediction is classified as a True Positive if the IoU is more than the IoU threshold and a False Positive is a prediction with an IoU below that threshold. If a detector or model fails to detect an object present in the ground truth, it is classified as a False Negative.

$$IOU = \frac{area\ of\ overlap}{area\ of\ union} =$$

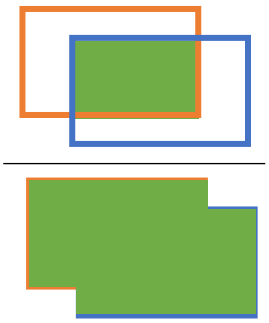

Fig. 8. Intersection over Union (IoU).

Moreover, each predicted bounding box outputs a confidence score that is used to measure the certainty of a prediction. A confidence threshold can be set to only return predictions with a confidence score at or higher than the threshold. The process of plotting a model's precision and recall as a function of the model's confidence threshold is the precision-recall curve. The precision-recall curve shows precision values for corresponding recall values for overall model performance plotted on a scale of 0 to 1 for both precision and recall. In a precision-recall curve, one plots recall on the *x*-axis and precision on the *y*-axis (Davis & Goadrich 2006).

Average precision (AP) is a popular metric for evaluating the accuracy of object detectors by estimating the area under the curve of the precision and recall relationship. The mean average precision (mAP) or mAP score is a common, singular evaluation metric that calculates the average precision (AP) for each class over all the confidence thresholds at a fixed IoU (Rezatofighi et al. 2019). Model performance in this study was reported using mAP as the single metric at an IoU of 0.5.

Table 5. Confusion matrix of 2x2 describing the binary classification with ground truth or actual values on one axis and predicted on another.

PREDICTION	GROUND TRUTH/ACTUAL: Negative	GROUND TRUTH/ACTUAL: Positive
Negative	TRUE NEGATIVE	FALSE NEGATIVE
Positive	FALSE POSITIVE	TRUE POSITIVE

$$precision = \frac{True\ Positive}{True\ Positive\ +\ False\ Positive} = \frac{True\ Positive}{All\ Predictions} \quad (3)$$

$$recall = \frac{True\ Positive}{True\ Positive\ +\ False\ Negative} = \frac{True\ Positive}{All\ Ground\ Truth} \quad (4)$$

J. Resources used

Table 6. List of resource computer specifications used to run object detection models.

GPU	GPU RAM (GB)	Source
Titan RTX	24	SDAV Lab
NVIDIA Tesla K40	32	Mana
NVIDIA Quadro RTX 5000	16	Mana

a. High Performance Computing

The high performance computing (HPC) cluster at the University of Hawai‘i, known as Mana, was used in this study to accelerate the model training and evaluation process for the transfer learning series of models. For each job, set to exclusive mode by default, 1 node (where each node has 22 GB of memory), 5 CPUs, 1 GPU, and 4 hours was requested.

K. Data analysis

Statistical analyses were performed using RStudio (RStudio Team 2020). Pearson's correlation was run to measure any statistical relationship between feature indices and model mAP. Four feature measurements (contrast, shape, size, and texture) were averaged across 50 samples per feature and repeated for all six targets of interest (guinea grass, miconia, red, brown, fine white, and skeleton). For all object detection models, each model instance was run five different times and mAP was averaged.

Model performance with transfer learning was compared to models without the addition of transfer learning by running 30 individual Student's t-tests with a Bonferroni p-value adjustment on each transfer learning instance (Abdi 2007). A Bonferroni correction was applied to avoid spurious significance by lowering the p-value to account for the number of comparisons being performed (Abdi 2007). When running 30 t-tests, one would expect some tests to reach the standard p-value cutoff of 0.05 due to chance because of the large number of tests run. The Bonferroni adjustment accounts for the large number of tests, helping strengthen the rigor of the results. A more stringent or smaller alpha level helps lower errors from rejecting the null hypothesis when it is true and reveal effects of transfer learning more thoroughly. Welch's t-test was also used to compare samples that did not meet the equal variance assumption (West 2021). The transfer learning mAP and the mAP of the target models without the transfer learning component were compared.

III. Results

A. Feature measurements

Feature measurements for all objects of interest are shown in Table 7. Object detection performance results for both source, target, and transfer learning instances are shown in Table 8 and 9 respectively. Regarding feature measurements, RMS contrast measurements were calculated with a range of 23.01 – 49.12. The circularity category was calculated with a range of 0.04 – 0.74 on a scale of 0-1. Measurements of size range between 624 – 38,608 square pixels. Texture measurements range between 334 – 2,102.67. The object of interest that measured the highest contrast, lowest circularity, and highest texture was skeleton. Guinea grass was measured with the lowest contrast score, second to lowest in circularity as well as size, and lowest in texture. Miconia was measured with the highest circularity and smallest size. Red, brown, and fine white yielded feature measurements that were similar in circularity although mid-range across all feature measurements.

Table 7. Species of interest and their corresponding feature measurements based on 50 samples.

Object of interest	RMS Contrast	Circularity	Size (sq. pixels)	Texture
Guinea grass	23.01 (± 4.285)	0.22 (± 0.137)	1151 (± 2,266.013)	334 (± 138.629)
Miconia	32.25 (± 8.273)	0.74 (± 0.079)	624 (± 152.189)	599 (± 356.664)
Red	23.85 (± 7.059)	0.41 (± 0.134)	38,608 (± 36,994.898)	454.04 (± 293.332)
Brown	33.75 (± 8.367)	0.49 (± 0.151)	18,932 (± 15,393.983)	751.94 (± 378.523)
Fine white	44.5 (± 9.952)	0.36 (± 0.159)	29,324 (± 27,509.082)	1,591.54 (± 746.656)
Skeleton	49.12 (± 7.514)	0.04 (± 0.024)	12,765 (± 10,672.891)	2,102.67 (± 710.849)

B. Object detection model performance without transfer learning

Object detection performance of source models and target models are shown in Table 8. Model performance of the source series that used 500 annotations in training demonstrated mAP that ranged between 0.024 – 0.669. Guinea grass had the lowest source model mAP while fine white had the highest mAP score. In the target series that used 25 annotations for training, mAP scores ranged between 0.063 – 0.407 where guinea grass had the lowest model performance and fine white resulted in the highest model performance. Nearly all objects of interest decreased in performance after the number of training annotations dropped from 500 to 25 in the source and target series except for guinea grass which had a slight increase in mAP.

Table 8. Object detection performance of both source series and target series. Source series was trained on 500 annotations and target series was trained with 25 annotations. Both series are measured by its mAP.

Object of Interest	Source mAP	Target mAP
Miconia	0.307 (± 0.062)	0.243 (± 0.095)
Guinea grass	0.024 (± 0.018)	0.063 (± 0.021)
Red	0.362 (± 0.041)	0.168 (± 0.029)
Brown	0.519 (± 0.024)	0.127 (± 0.010)
Fine white	0.669 (± 0.020)	0.407 (± 0.017)
Skeleton	0.576 (± 0.010)	0.240 (± 0.021)

C. Feature measurements & model performance correlations

The four correlation analyses performed between all source model mAP and individual feature measurements resulted in a significant correlation among the contrast measure and source model mAP ($r(4) = 0.82$, $p = 0.045$). A strong correlation was found between texture and source model mAP ($r(4) = 0.77$, $p = 0.073$), weak correlation between source mAP and circularity measures ($r(4) = –0.096$, $p = 0.86$) as well as moderate correlation for the size measure ($r(4) = 0.54$, $p = 0.27$) (Figure 9). I found there was no significance in correlations between source model mAP and circularity, size, and texture.

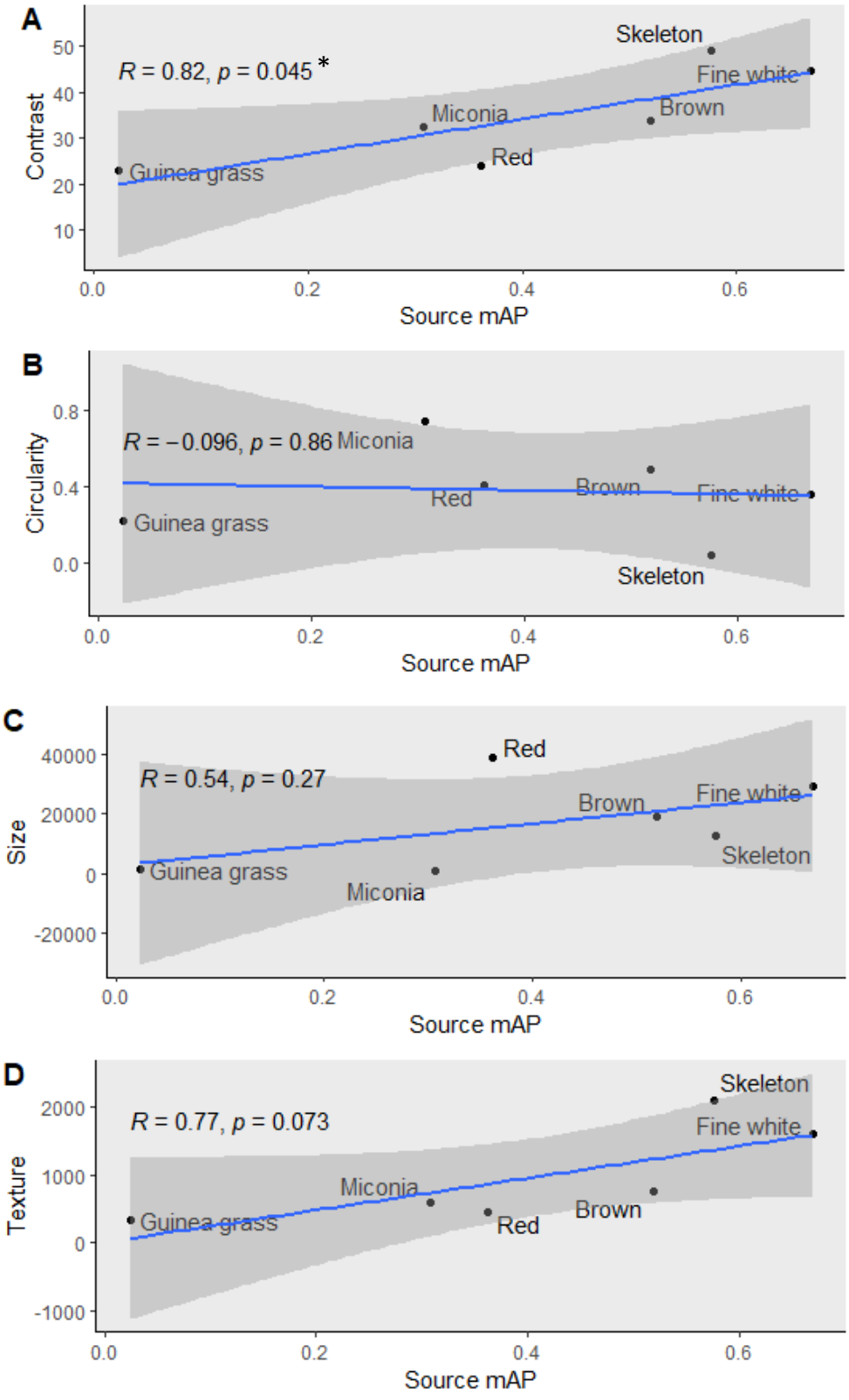
A
Contrast
R = 0.82, p = 0.045 *
Skeleton
Fine white
Brown
Miconia
Red
Guinea grass
Source mAP
B
Circularity
R = −0.096, p = 0.86
Miconia
Red
Brown
Fine white
Guinea grass
Skeleton
Source mAP
C
Size
R = 0.54, p = 0.27
Red
Brown
Fine white
Skeleton
Guinea grass
Miconia
Source mAP
D
Texture
R = 0.77, p = 0.073
Skeleton
Fine white
Miconia
Brown
Guinea grass
Red
Source mAP

Fig. 9. Correlation graphs of the four feature measurements and its corresponding source mAP score per target of interest. The significant correlation between source mAP and contrast is identified by an asterisk (*).

D. Performance and comparison of transfer learning models

Thirty t-tests were conducted between each target and transfer learning model's mAP scores with a Bonferroni-adjusted alpha level of 0.00167 per test (0.05/30). 12 out of 30 t-tests were significantly different, indicating there were significant improvements or decreases in transfer learning performance. Out of the 12 significant t-tests, 9 transfer learning instances had higher mAP scores than the target mAP score (Table 9).

Table 9. Object detection performance measured by the difference in mean average precision between transfer learning and target mAP from Table 8. A positive value indicates the Transfer learning model had a higher mAP score than the Target model. Transfer learning models where mAP increased by at least 0.1 are green bolded and contain an asterisk (*). All transfer learning models that were found significantly different than the target mAP only displays an asterisk (*).

Target species

Source species	**Guinea grass**	**Miconia**	**Red**	**Brown**	**Fine white**	**Skeleton**
Guinea grass	/	-0.028 (± 0.106)	-0.070 (± 0.031)	-0.019 (± 0.017)	-0.098* (± 0.035)	0.019 (± 0.010)
Miconia	-0.026 (± 0.020)	/	0.014 (± 0.033)	-0.117* (± 0.008)	-0.086* (± 0.026)	0.010 (± 0.016)
Red	-0.039 (± 0.016)	-0.145 (± 0.098)	/	**0.242*** (± 0.012)	-0.040 (± 0.059)	0.069* (± 0.030)

Brown	-0.040 (± 0.021)	-0.162 (± 0.101)	**0.232*** (± 0.033)	/	-0.001 (± 0.020)	0.068* (± 0.037)
Fine white	0.092 (± 0.043)	-0.096 (± 0.141)	**0.191*** (± 0.031)	**0.267*** (± 0.023)	/	**0.230*** (± 0.026)
Skeleton	**0.138*** (± 0.020)	-0.198 (± 0.092)	**0.189*** (± 0.025)	0.018 (± 0.010)	-0.010 (± 0.042)	/

IV. Discussion

Transfer learning is commonly applied where a domain task is pre-trained on an existing, large-scale datasets such as ImageNet, to provide a baseline for the object detection network's new task. In this study, models were first pre-trained using the ImageNet dataset, but underwent a second transfer learning process by using pre-trained networks from individualized plant models trained with significantly less data than available in the ImageNet dataset. As the ImageNet database contains thousands of everyday objects like "dog", "balloon", or "car", this was done to explore any transfer learning benefits of leveraging plant-only targets that are similar to one another versus when they are different. Apart from a select few models, the resulting mAP performance scores reflect poor overall detection (Table 8 & 9), likely due to limited training examples or potentially insufficient training time. Although mAP scores were low, the main objectives of this study were concentrated on the increase or decreases in improvement between target model mAP and transfer learning mAP as well as determining how features impact detection performance.

From our series of model runs and data collection, transfer learning was found to be significantly beneficial for some sources and targets that were similar. However, transfer learning is not a silver bullet and should be used with discretion. This study aimed to compare

performance between transfer learning from ImageNet alone, a dataset containing a wide variety of objects and image types, and a constrained dataset of only plants and aerial imagery. Out of 30 transfer learning models ran on the constrained dataset, majority performed worse or negligibly the same (21/30 models) and some performed significantly better (9/30 models) than transfer learning models using ImageNet only (Table 9). This study found transfer learning from a constrained dataset (single object category of the same data type) still requires additional similarities for successful transfer learning. Based on these results, choosing which dataset to transfer learn from should be selectively considered in applications of object detection that undergo a second level of transfer learning. Although all targets of interest in this study were plants, using three different types of plants (grass, shrub, and tree) helped reveal changes or improvements when the source and target were more similar versus when they were very different.

The significant increases in mAP performance for transfer learning models come from source and targets that are more similar while significant decreases in mAP are between source and targets with dissimilar features (Table 9). Similar objects of interest in this study can be defined as objects that, when grouped, do not have significant differences across all feature measures of contrast, circularity, size, and texture (Figure 10 & 11). As a group, red, brown, and fine white share no significant differences across all contrast, circularity, size, and texture measurements ($p > 0.05$). In terms of the size measure, when all ROD classes (red, brown, fine white, and skeleton) are grouped they share no difference in size ($p < 0.1$). However, when miconia and all ROD classes are grouped, there is a significant difference in size ($p < 0.0001$). When miconia and guinea grass are grouped there is no difference in size ($p > 0.1$). Further "groups" are revealed in circularity where miconia, red, brown, and fine white classes are found to be similar ($p > 0.5$). In addition, red, brown, and fine white resulted in no difference in texture as well as contrast ($p = 1$).

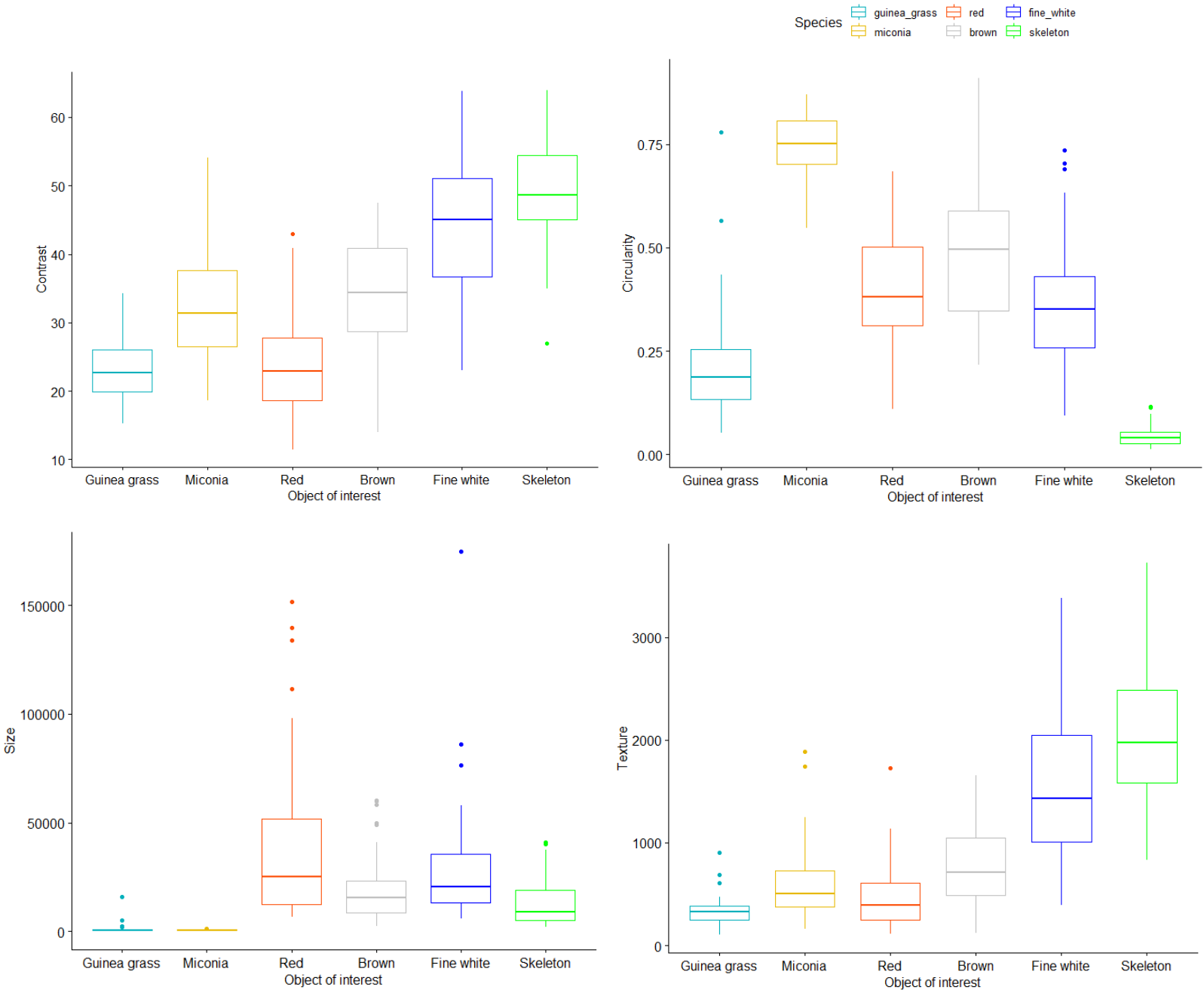

Fig. 10. Boxplot visualization of each feature measurement's distribution (circularity, size, texture, and contrast) for all six objects of interest.

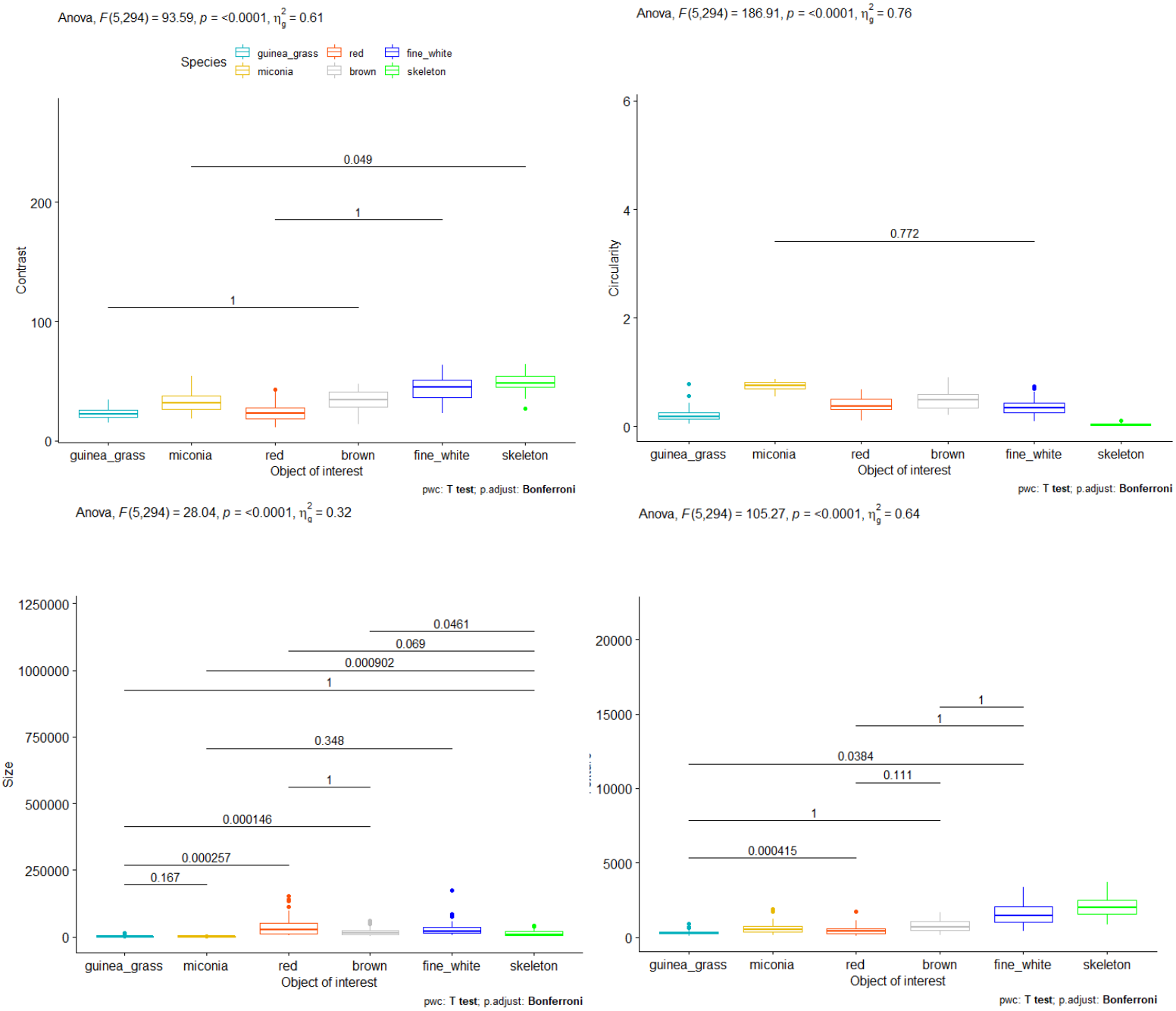

Fig. 11. The same boxplot visualization seen in Figure 10 with the addition of *p*-values from pairwise t-tests. Only *p*-values greater than 0.0001 are included to emphasize nuanced patterns of similarity or difference when a variety of species combinations are grouped together.

The most improvement occurred when source and target objects were of ROD classes (red, brown, fine white, and skeleton), suggesting transferability may be most effective when objects of interest exhibit similarities in multiple feature categories. Here it is important to note that when some ōhi‘a trees are in transition to its following symptomatic stage, the distinction between classes can be blurred, introducing occasional overlap across ROD class datasets. Effects of overlap may also explain transfer learning improvements when ROD classes were involved. The significant decreases in transfer learning models of miconia to brown, miconia to fine white, and guinea grass to fine white suggest these plants may not similar enough to be applied in a transfer learning situation effectively. A study focused on identifying maize in UAV aerial imagery compared the performance of a model pretrained on ImageNet versus that of a model pretrained on plant images and found that the model pretrained on ImageNet did worse than the model pretrained on plant images (Cai et al. 2020). They concluded that the source domain is too different than the target domain (the more general ImageNet vs. UAV plant images) (Cai et al. 2020). The variation in results after transfer learning support that different types of knowledge may transfer better or worse depending on the task similarity where low-level information may transfer across closely related tasks while high-level concepts may transfer across pairs of less similar tasks (Taylor & Stone 2009). Although our study’s workflow does not follow the exact framework, the results of our study corroborate intuitions that more similar plants transfer more effectively than different domains (Ghazi et al. 2017). New image datasets, such as the NEON Crowns dataset, have emerged that include tree crowns rather than everyday objects and could be potentially incorporated into the current machine learning pipeline to improve detection of invasive plants in Hawai‘i (Weinstein et al. 2021).

More evidence of shared features between ROD subclasses can be seen in Figure 12 where brown, fine white, and skeleton share the most overlap when size, texture, and contrast variables are plotted in 3D. Contrary to my hypothesis that predicted objects with higher shape irregularity, such as guinea grass, would produce higher model performance, circularity was not correlated with mAP in the source, target, and transfer learning model series. This may be due to the limitations of the circularity measure being unable to fully convey the unique shapes for each target of interest. I also hypothesized guinea grass would improve scores most when applied in transfer learning, but “skeleton” exhibited the lowest circularity, highest contrast, and was the only class to improve guinea grass’ mAP in the transfer learning instances. Guinea grass’ source

model mAP was lowest while skeleton's source model mAP was second highest. Transfer learning from skeleton to guinea grass resulted in improvement. However, transfer learning from guinea grass to skeleton resulted in a performance decrease. In addition to the intuition that transfer learning is likely most effective when source and target domains are similar (Cai et al. 2020), transfer learning success may also be due to performance of the source model: skeleton's source model mAP was fairly high whereas guinea grass' source model mAP was much lower. Moreover, guinea grass' low detection performance may be a product of the blurry imagery used to train guinea grass during an early phase of DSLR camera aerial imagery collection before protocols and camera settings had been optimized.

Although guinea grass had the lowest performance in the source, target, and transfer learning model series, there are feasible improvements that can be made to potentially increase its model performance. In future data collection surveys for guinea grass, it may be worthwhile to capture imagery closer to ground allowing edges of guinea grass to be more distinguishable. In a detection study focused on grass and grass-like weeds, researchers discussed and recommended taking imagery at a closer level as the target grasses and non-target surroundings can be indistinct or of similar kinds of grasses (Bullock et al. 2019). Additionally, the influence of growth stages could be taken advantage of when collecting data for guinea grass. Based on the significance of contrast features, data for guinea grass could be collected during seasons that allow guinea grass to appear more pronounced against its surroundings. Collecting higher quality data for guinea grass should be implemented into new guinea grass detection models to determine if detection improves before applying transfer learning from guinea grass models to other grasses. Similarly, as transfer learning between ROD classes produced improvements in this study, I suspect analogous results when transfer learning from guinea grass to other grass species. However, it is important to note that grasses have been challenging targets to detect because of their indistinct edges against backgrounds of similar plants or grasses in their surroundings (Bullock et al. 2019; Wu et al. 2021)

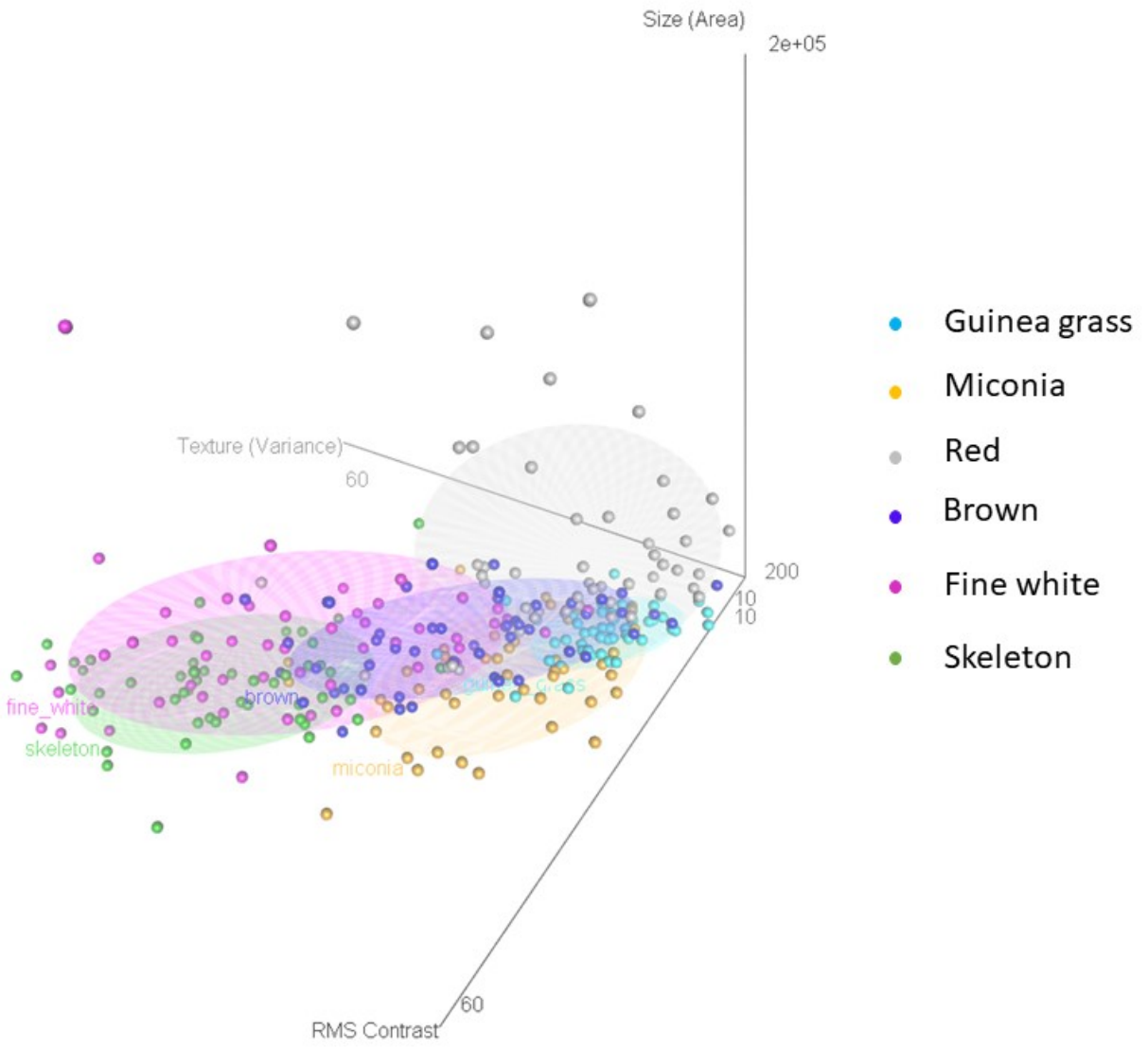

Fig. 12. 3D scatterplot of guinea grass, miconia, red, brown, fine white, and skeleton feature measures of size, texture, and RMS contrast. Each class is grouped in its own ellipses.

This study contains limitations of model processing time that restrained our ability to run more than five model runs per source, target, and transfer learning instance. Moreover, small datasets per target of interest constrain the distribution of feature measurements, ultimately producing a confounding factor for concluding definitive feature measurements. Thus, additional model performance data and a greater distribution of data per targets of interest would be required to gather more conclusive inferences based on feature measurements. However, in a scenario where a greater number of annotations were available for each target of interest's source model, the expected model performance after transfer learning may not necessarily be higher

with more source annotations. A greater number of annotations per target of interest is suspected to allow models to generalize better, but it is more likely that the dominating factor contributing to transfer learning success depends on the degree of similarity between the source and target objects.

For future work, another invasive plant of particular interest is faya tree (*Morella faya*). *Morella faya* is another highly invasive and nitrogen-fixing tree found in Hawai‘i that originates from the Canary Islands, Madeira, and Azores (Wagner et al. 1999). Specifically, faya tree is widely found in Hawai‘i Volcanoes National Park and is considered a non-native weed to be controlled within park boundaries (Benitez et al. 2012). Faya tree presents another study case in which limited aerial imagery data is currently available. If faya tree were to be considered as the next species of interest for object detection, I generated a set of preliminary feature measurements (n=10) from imagery collected within HAVO Park boundaries to demonstrate how these may guide how transfer learning should be applied (Table 10). In the annotated faya tree data, contrast measurements were found closest to fine white's contrast. However, this may potentially be largely due to the data containing many shadows from trees as this imagery was collected early morning. Circularity measure for faya tree was between the range of miconia and red's circularity measure. The size of faya tree was within range of all ROD classes. Faya tree texture is most like brown's texture measurement. Faya tree's feature measurements and all ROD class feature measurements appeared to be most alike. When comparing similarity in feature measurements and the performance of red, brown, fine white, and skeleton models that resulted in improvements from this study, transfer learning from fine white or brown as the source model may benefit faya tree as a target task rather than guinea grass, miconia, red, or skeleton. As contrast and texture measurements were found to be highly correlated with source model performance, faya tree shares a similar contrast measure with fine white and demonstrates a similar texture measure with brown. Overall, because faya tree measures similarly to multiple feature measurements of ROD classes, I suspect faya tree would be a good target candidate to use in transfer learning with specifically fine white or brown as the source model.

Table 10. Faya tree and its corresponding feature measurements.

Object of interest	RMS Contrast	Circularity	Size (sq. pixels)	Texture
Faya tree	45.03 (± 2.984)	0.62 (± 0.091)	44,551 (±20,735.612)	736.08 (± 235.542)

While transfer learning can be an approach to leverage existing data, it is also an opportunity to apply transfer learning for detection of incipient invasive species. Early-detection of invasive plants is an essential component for land management to identify and eradicate new infestations of nonnative plants (U.S. Geological Survey 2012). Populations of new nonnative species are usually small, resulting in low amounts of collectable data. Approaches of transfer learning can capitalize on existing models to help resolve data limitations of incipient species. As object detection gains popularity in monitoring and controlling invasive plants, transfer learning can be used as a tool for land managers and early-detection of invasive species programs. Moreover, object detection of plants with sufficient data should be expanded to a variety of plant species. Robust models from a diverse suite of plants can be used as effective source models for transfer learning. Because it is likely that the incipient species will be random, a suite of source models allows transfer learning from a greater pool of targets that may share similarities.

In addition, contrast was the only image characteristic in this study that was significantly correlated with source model mAP. In considering other species of interest where object detection and transfer learning could be applied in Hawai‘i, there is evidence that when source plant and target plant are more similar the more likely transfer learning will be effective. Imagery of species of interest with median contrast measurements were found to have an impact on mAP. Other species candidates in which transfer learning could be applied should consider the similarity and image characteristics of the source and target plants. The benefit of transfer learning should be further experimented using different species than those used in this study to better understand which data transfer learning can be leveraged best when plants found in Hawai‘i are the objects of interest. Additionally, this study focused on single class detection, but

experimenting with models trained on multiple "source" classes may reveal other advantages to transfer learning.

V. Conclusions

Hawaiian landscapes are highly impacted by invasive plants that alter the structure and function of native ecosystems and exacerbate native habitat degradation, suppress ecosystem resilience, and diminish endemic diversity in Hawai‘i (Vorsino et al. 2014). Efforts to routinely map, monitor, and detect emerging invasive species commonly use high-resolution aerial imagery and human observers for identification (Rodriguez et al. 2021). As we transition from the use of human observers to a more automated approach for invasive species target detection, Hawai‘i may serve as a model system through innovative approaches to reduce the impacts of invasive species (Cordell 2021). Deep learning shows promise as a tool in which all land management agencies throughout Hawai‘i can integrate alongside or to reduce efforts from the numerous strategies required for invasive species control and early detection.

While large volumes of training examples are considered vital for successful applications of deep learning, examples are challenging to collect and acquire for detecting new invasive plant species. To address scenarios where copious training data is not readily available, a transfer learning component was incorporated in this study to explore the transferability of a suite of plants to one another to determine how much different domains affected transfer learning performance. By understanding how model performance behavior can be best leveraged across various plants of interest, it may be possible that data usage and collection could be significantly maximized.

In this thesis, I demonstrate the outcomes of transfer learning when applied to object detection of select invasive plants on Big Island. I also present feature measurements per invasive plant to better understand how image features impact model performance. Transfer learning improved most when the source or target involved red, brown, fine white, and skeleton ROD trees. Transfer learning also significantly improved when the skeleton source model was applied to the guinea grass target model. Source and target plants that were more similar resulted in the highest increases in mAP. Out of circularity, contrast, size, and texture, contrast was most

correlated with model performance. Implementation of transfer learning in this study for object detection of invasive plants has shown to be effective even when data is greatly reduced. The method of object detection in this study can be easily adapted to experiment with different plants of interest on and outside of Hawai‘i. Future research should integrate more source and target plants that are alike to further confirm the effects of transfer learning. As this study used source and target domains that were both plants in aerial imagery, transfer learning improvements mainly occurred between plants that were very similar in image contrast and texture.

For land management in Hawai‘i and elsewhere, object detection can be used as a tool to help detect arguably any object found in an image. Some of the biggest challenges in using these methods is acquiring sufficient data and its relatively new application in the ecological space. In scenarios where there may not be enough data to apply object detection for a specific incipient invasive plant, for example, land managers can consider incorporating transfer learning. Here, land managers can decide to leverage existing data or detection model of another similar plant (source model) to apply towards creating a new detection model of the incipient invasive plant (target model). To apply transfer learning, a readily available and well performing detection model is needed to use as a starting point to train the incipient species of interest. This study emphasizes it is important that land managers use discretion when considering which species' detection model to involve in the transfer learning process. Based on the findings of this study, selecting a source model whose plant type is the same plant as the target plant of interest may produce greater transfer learning success. For example, transferring information from a tree model to help train another tree model performed better than transfer learning from a grass model to a tree model. In addition to selecting the same kind of plant, choosing a source whose colors show up prominently against its surrounding background (such as fine white ROD and skeleton ROD in this study) may produce the highest performing models and contribute towards improvements from transfer learning. Moreover, choosing a source model that already performs well also helps increase transfer learning success. Outcomes of this study should be further investigated and expanded to different plants of interest to better confirm which source and target species provide the best improvements after transfer learning.

This study applied transfer learning from plant aerial images to another set of plant aerial images. The results indicate that transfer learning success between plants require greater

similarity between source and target objects in addition to being a single object category of the same data type. It is generally thought that transfer learning should always help as long as there is some similarity between source and target, but this thesis shows that it is not always the case. Overall, what was found in this study can further advance the knowledge base of deep learning tools used to detect invasive species for natural resource management. Focusing on improvements or ways to make these detection systems more efficient may support federal, international, and local initiatives to control established harmful invasive species infestations or prevent incipient and potentially harmful invasions at earlier stages.

References

Abadi, M., Agarwal A., Barham P., Brevdo E., Chen Z., Citro C., Corrado G.S., David A., Dean J., Devin M., Ghemawat S., Goodfellow I., Harp A., Irving G., Isard M., Jia Y., Jozefowicz R., Kaiser L., Kudlur M., Levenberg J., Mane D., Monga R., Moore S., Murray D., Olah C., Schuster M., Shlens J., Steiner B., Sutskever I., Talwar K., Tucker P., Vanhoucke V., Vasudevan V., Viegas F., Vinyals O., Warden P., Wattenberg M., Wicke M., Yu Y., X. Zheng. 2016. TensorFlow: large-scale machine learning on heterogeneous distributed systems. Computing Research Repository (CoRR) **1603**: 1-19.

Abdi, H. 2007. The Bonferroni and Sidak corrections for multiple comparisons. Salkind NJ (ed) Encyclopedia of measurement and statistics.

Adam, E. V., B. F. Lucas, A. A. Fred, E. M. Stephen, D. J. James, P. P. Jonathan, Sam 'Ohukani'Ohi'a Gon, and A. K. Gregory. 2014. Modeling Hawaiian ecosystem degradation due to invasive plants under current and future climates. PLoS ONE **9**(5)**:**e95427.

Ahmad, M., Abdullah, M., & Han, D. 2020. Small object detection in aerial imagery using RetinaNet with Anchor Optimization. 2020 International Conference on Electronics, Information, and Communication (ICEIC).

Anderson CB. 2018. Biodiversity monitoring, earth observations and the ecology of scale. Ecology Letters **21**:1572–1585.

Asner GP, Martin RE, Knapp DE, Kennedy-Bowdoin T. 2009. Effects of Morella faya tree invasion on aboveground carbon storage in Hawai‘i. Biological Invasions **12**:477–494.

Babatunde OH, Armstrong L, Leng J, Diepeveen D. 2015. A survey of computer-based vision systems for automatic identification of plant species. Journal of Agricultural Informatics **6**.

Ball JE, Anderson DT, Chan CS. 2017. Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. Journal of Applied Remote Sensing **11**:1.

Barnes I, Fourie A, Wingfield MJ, Harrington TC, McNew DL, Sugiyama LS, Luiz BC, Heller WP, Keith LM. 2018. New *Ceratocystis* species associated with rapid death of *Metrosideros polymorpha* in Hawai`i. Persoonia. **40**:154-181.

Benitez, D.M., R. Loh, T. Tunison, N.G. Zimmer, J. Makaike, R. Mattos and M. Casali. 2012.The distribution of invasive plant species of concern in the Kīlauea and Mauna Loa strip areas of Hawai‘i Volcanoes National Park, 2000-2010. Technical Report No. 179. The Hawai`i-Pacific Islands Cooperative Ecosystem Studies Unit & Pacific Cooperative Studies Unit, University of Hawai`i, Honolulu, Hawai`i. 120 pp.

Bhuiyan A.A., Khan A.R. 2018. Image Quality Assessment Employing RMS Contrast and Histogram Similarity. The International Arab Journal of Information Technology. **15**:983-989.

Blaschke, T. Hay G.J., Kelly M., Lang S., Hofmann P., Addink E., Queiroz Feitosa R., Van der Meer F., Van Der Werff H., Van Coillie F., Tiede D. 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. ISPRS Journal of Photogrammetry and Remote Sensing **87**:180-191.

Borenstein E., Malik J. Shape guided object segmentation. 2006. Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. **1**: 969-976.

Boroweic ML., Frandsen P., Dikow R., McKeenan A., Valentini G., White AE. 2021. Deep learning as a tool for ecology and evolution. EcoEvoRxiv. 26.

Brodrick PG, Davies AB, Asner GP. 2019. Uncovering Ecological Patterns with Convolutional Neural Networks. Trends in Ecology & Evolution **34**:734–745.

Bullock D, Mangeni A, Wiesner-Hanks T, DeChant C, Stewart EL, Kaczmar N, Kolkman JM, Nelson RJ, Gore MA, Lipson H. 2019. Automated weed detection in aerial imagery with context. Computer Vision and Pattern Recognition.

Cai, E., Baireddy S., Yang C., Crawford M., Delp, E.J. 2020. Deep transfer learning for plant center localization. Computer Vision and Pattern Recognition. **1**: 277-284.

Chollet F. 2015. http://github.com/keras-team/keras

Cordell S. 2021. Regional Summaries: Hawaii and US-Affiliated Pacific Islands. Pages 343–352 in T.M. Poland; T. Patel-Weynand; D.M. Finch; C. Ford Miniat; D.C. Hayes; V.M.

Lopez, eds. Invasive Species in Forests and Rangelands of the United States: A Comprehensive Science Synthesis for the United States Forest Sector. Heidelberg, Germany: Springer International Publishing.

Culman M, Delalieux S, Tricht KV. 2020. Palm tree inventory from aerial images using RetinaNet. Proceedings of the 2020 Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS). 314-317.

Csillik O, Cherbini J, Johnson R, Lyons A, Kelly M. 2018. Identification of Citrus Trees from Unmanned Aerial Vehicle Imagery Using Convolutional Neural Networks. Drones **2**:39.

Davis J, Goadrich M. 2006. The relationship between precision-recall and ROC curves. In Proceedings of the 23rd International Conference on Machine Learning.

Day O, Khoshgoftaar TM. 2017. A survey on heterogeneous transfer learning. Journal of Big Data. 4(**29**).

Dhillon, A., and G. Verma. 2019. Convolutional neural network: a review of models, methodologies and applications to object detection. Progress in Artificial Intelligence **9**:85-112.

dos Santos AA, Marcato Junior J. Araújo MS, Di Martini DR, Tetila EC, Siqueira HL, Aoki C, Eltner A, Matsubara ET, Pistori H, et al. 2019. Assessment of CNN-based methods for individual tree detection on images captured by RGB cameras attached to UAVs. Sensors. **19**: 3595.

Ellen, J.S., Graff, C.A., Ohman, M.D. 2019. Improvement plankton image classification using context metadata. Limnology and Oceanography: Methods. 17(**8**): 439-461.

Exelis Visual Information Studios. 2010. Texture Metrics Background. Boulder, Colorado. Exelis Visual Information Solutions.

Fortini LB, Kaiser LR, Keith LM, Price J, Hughes FR, Jacobi JD, Friday JB. 2019. The evolving threat of Rapid ʻŌhiʻa Death (ROD) to Hawaiʻi's native ecosystems and rare plant species. Forest Ecology and Management. **448**: 376-385.

Fuentes-Pacheco J, Torres-Olivares J, Roman-Rangel E, Cervantes S, Juarez-Lopez P, Hermosillo-Valadez J, Rendón-Mancha JM. 2019. Fig Plant Segmentation from Aerial Images Using a Deep Convolutional Encoder-Decoder Network. Remote Sensing **11**:1157.

Fung Associates, Incorporated, and SWCA Environmental Consultants. 2019. Natural resource condition assessment: Hawai`I Volcanoes National Park. Natural Resource Report NPS/HAVO/NRR – 2019/1967. National Park Service, Fort Collins, Colorado.

Frazor, R., and W. Geisler. 2006. Local luminance and contrast in natural images. Vision Research **46**:1585-1598.

Gaiser, H. de Vries, M. Williamson, A. Hennon, Y. Lacatusu, V. Vidosits, A. Gratie, C. et al. 2021. keras-retinanet. Version 0.5.1. Available from https://github.com/fizyr/keras-retinanet.

Ghazi, M.M., Y, Berrin., E, Aptoula. 2017. Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing. **235**: 228-235.

González-Muñoz N, Bellard C, Leclerc C, Meyer J-Y, Courchamp F. 2015. Assessing current and future risks of invasion by the "green cancer" Miconia calvescens. Biological Invasions **17**:3337–3350.

Gopalakrishnan K, Khaitan SK, Choudhary A, Agrawal A. 2017. Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection. Construction and Building Materials **157**:322–330.

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. 2016. Deep learning for visual understanding: A review. Neurocomputing. **187**:127–48.

Gould, S., Gao T., Koller D. 2009. Region-based segmentation and object detection. Advances in Neural Information Processing Systems. **22**: 1-9.

Jamil N, Hussin NAC, Nordin S, Awang K. 2015. Automatic Plant Identification: Is Shape the Key Feature? Procedia Computer Science **76**:436–442.

Joshi KA, Mulder RA, Rowe KMC. 2017. Comparing manual and automated species recognition in the detection of four common south-east Australian forest birds from digital field recordings. Emu - Austral Ornithology **117**:233–246.

Hallé F., Oldeman R.A.A., Tomlinson P.B. 1978. Tropical trees and forests: an architectural analysis. Springer-Verlag.

Haralick RM, Shanmugam K, Dinstein I. 1973. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics **SMC-3**:610–621.Ho C-Y, Tsai M-Y, Huang Y-L, Kao W-Y. 2015. Ecophysiological factors contributing to the invasion of Panicum maximum into native Miscanthus sinensis grassland in Taiwan. Weed Research **56**:69–77.

Hong S-J, Han Y, Kim S-Y, Lee A-Y, Kim G. 2019. Application of Deep-Learning Methods to Bird Detection Using Unmanned Aerial Vehicle Imagery. Sensors **19**:1651.

Horwath JP, Zakharov DN, Mégret R, Stach EA. 2020. Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images. Npj Computational Materials. **6**:108.

Kadir A., Nugroho L.E., Susanto A., Santosa P.I. 2011. A comparative experiment of several shape methods in recognizing plants. International Journal of Computer Science & Information Technology (IJCSIT) **3**: 256-263.

Kattenborn, T., Leitloff, J., Schiefer, F. and Hinz, S., 2021. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing **173**: pp.24-49.

Klinken RDV, Friedel MH. 2017. Unassisted invasions: understanding and responding to Australia's high-impact environmental grass weeds. Australian Journal of Botany **65**:678.

Kornblith S., Shlens J., Q.V. Le. 2019. Do better ImageNet models transfer better? The IEEE Conference on Computer Vision and Pattern Recognition (CVPR): 2661-2671.

Krizhevsky A, Sutskever I, Hinton GE. 2017. ImageNet classification with deep convolutional neural networks. Communications of the ACM **60**:84–90.

LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. Nature **521**:436–444.

Leary J, Mahnken B, Wada C, Burnett K. 2018. Interpreting life-history traits of miconia (Miconia calvascens) through management over space and time in the east Maui watershed, Hawai'i (USA). Invasive Plant Science and Management **11**:191-200.

Li, K., G. Wan, G. Cheng, L. Meng, and J. Han. 2020. Object detection in optical remote sensing images: A survey and a new benchmark. ISPRS Journal of Photogrammetry and Remote Sensing **159**:296-307.

Li W, Fu H, Yu L, Cracknell A. 2016. Deep Learning Based Oil Palm Tree Detection and Counting for High-Resolution Remote Sensing Images. Remote Sensing **9**:22.

Lin T-Y, Goyal P, Girshick R, He K, Dollar P. 2018. Focal loss for dense object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence **1**:1–10.

Lin, T., P. Goyal, R. Girshick, K. He, and P. Dollar. 2020. Focal Loss for Dense Object Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 42:318-327.

Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C, C -Y, Berg AC. 2016. Single Shot MultiBox detector. In Proceedings of the Computer Vision – ECCV 2016. 21-37.

Loope, L., Hughes, F., Keith, L., Harrington, T., Hauff, R., Friday, J., Ewing, C., Bennett, G., Cannon, P., Atkinson, C., Martin, C., Melzer, M. 2016. Guidance document for Rapid Ohia Death: background for the 2017-2019 Rod Strategic Response Plan.

Lowe S., Brown M., Boudjelas S., De Porter M. 2000. 100 of the world's worst invasive alien species: a selection from the Global Invasive Species Database. The Invasive Species Specialist Group (ISSG) 12 pp.

Lu H, Fu X, Liu C, Li L-G, He Y-X, Li N-W. 2017. Cultivated land information extraction in UAV imagery based on deep convolutional neural network and transfer learning. Journal of Mountain Science **14**:731–741.

MacIntyre B, Cowan WB. 1994. A practical approach to calculating luminance contrast on a CRT. ACM Transactions on Graphics **11**:336–347.

Ma L, Liu Y, Zhang X, Ye Y, Yin G, Johnson BA. 2019. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing **152**: 166–177.

MacNeil, L., Missan Sg., Luo J., Trappenberg, T., LaRoche J. 2021. Plankton classification with high-throughput submersible holographic microscopy and transfer learning. BMC Ecology and Evolution. **21**: 123.

Mader S, Grenzdörffer GJ. 2016. Automatic Sea Bird Detection From High Resolution Aerial Imagery. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XLI-B7**:299–303.

Maggiori E, Tarabalka Y, Charpiat G, Alliez P (2017) Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification. IEEE Transactions on Geoscience and Remote Sensing **55**:645–657.

Mantoani MC, Torezan JMD. 2016. Regeneration response of Brazilian Atlantic Forest woody species to four years of Megathyrsus maximus removal. Forest Ecology and Management **359**:141–146.

Marmanis D, Datcu M, Esch T, Stilla U. 2016. Deep Learning Earth Observation Classification Using ImageNet Pretrained Networks. IEEE Geoscience and Remote Sensing Letters **13**:105–109.

Meyer J.Y., Loope L., A.C. Goarant. 2011. Strategy to control the invasive alien tree Miconia calvascens in Pacific islands: eradication, containment or something else? Islands invasives: eradication and management :92-96.

Miao Z., Gaynor K.M., Wang J., Liu Z., Muellerklein O., Norouzzadeh M.S., McInturff A., Bowie R.C.K., Nathan R., Yu S.X., Getz W.M. 2019. Insights and approaches using deep learning to classify wildlife. Scientific Reports. **9**:1-9.

Morgan JL, Gergel SE, Coops NC. 2010. Aerial Photography: A Rapidly Evolving Tool for Ecological Management. BioScience **60**:47–59.

Mortenson LA, Hughes RF, Friday JB, Keith LM, Barbosa JM, Friday NJ, Liu Z, Sowards TG. 2016. Assessing spatial distribution, stand impacts and rate of Ceratocystis fimbriata induced ʻōhiʻa (Metrosideros polymorpha) mortality in a tropical wet forest, Hawaiʻi Island, USA. Forest Ecology and Management **377**:83–92.

Moulden, B., F. Kingdom, and L. Gatley. 1990. The Standard Deviation of Luminance as a Metric for Contrast in Random-Dot Images. Perception **19**:79-101.

Muthuraja M, Arriaga O, Ploger P, Kirchner F, Valdenegro-Toro M. 2020. Black-box optimization of object detector scales. Computer Vision and Pattern Recognition.

Ochoa KS, Guo Z. 2019. A framework for the management of agricultural resources with automated aerial imagery detection. Computers and Electronics in Agriculture **162**:53–69.

Oghaz MMD, Razaak M, Kerdegari H, et al (2019) Scene and Environment Monitoring Using Aerial Imagery and Deep Learning. 2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS).

Olson E. 2011. Particle shape factors and their use in image analysis - part I: theory. Journal of GXP Compliance. **15**: 85-96

Padilla, R., Netto, S.L., da Silva, E.A.B. 2020. A survey on performance metrics for object-detection algorithms. 2020 International Conference on Systems, Signals and Image Processing (IWSSIP).

Perez-Sanz F, Navarro PJ, Egea-Cortines M. 2017. Plant phenomics: an overview of image acquisition technologies and image data analysis algorithms. GigaScience **6:**1-18.

Perroy, R., T. Sullivan, D. Benitez, R. Hughes, L. Keith, E. Brill, K. Kissinger, and D. Duda. 2021. Spatial Patterns of ʻŌhiʻa Mortality Associated with Rapid ʻŌhiʻa Death and Ungulate Presence. Forests 12:1035.

Peters DPC, Havstad KM, Cushing J, Tweedie C, Fuentes O, Villanueva-Rosales N. 2014. Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology. Ecosphere **5:**1-15.

Pitts W., McCulloch W.S. 1947. How we know universals the perception of auditory and visual forms. The Bulletin of Mathematical Biophysics. **9**(3): 127-147.

Rahimi-Nasrabadi, H., J. Jin, R. Mazade, C. Pons, S. Najafian, and J. Alonso. 2021. Image luminance changes contrast sensitivity in visual cortex. Cell Reports **34**:108692.

Rawat W, Wang Z. 2017. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. Neural Computation **29**:2352–2449.

Redmon J, Divvala S, Girshick R, Farhadi A. You Only Look Once: unified, real-time object detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 779-788.

Rezatofighi H, Tsoi N, Gwak J, Sadeghian A, Reid I, Savarese S. 2019. Generalized Intersection over Union: a metric and a loss for bounding box regression. Computer Vision and Pattern Recognition.

Rodriguez R III, Perroy RL, Leary J, Jenkins D, Panoff M, Mandel T, Perez P. Comparing interpretation of high-resolution aerial imagery by humans and artificial intelligence to detect an invasive tree species. Remote Sensing. 2021. **13**:3503.

RStudio Team. 2020. RStudio: integrated development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Russ J.C., Russ C. How not to be a circle: dimensionless ratios as shape descriptors. 2017. European Light Microscopy Initiative. **48**: 59-72.

Seals, M. 2019. On the robustness of object detection based deep learning models. Master's Thesis: University of Tennessee.

Sestili C. 2018. Deep learning: going deeper toward meaningful patterns in complex data. Carnegie Mellon University, Software Engineering Institute Blog. Available from https://insights.sei.cmu.edu/blog/deep-learning-going-deeper-toward-meaningful-patterns-in-complex-data/ (Accessed November 29, 2021).

Scott GJ, England MR, Starms WA, Marcum RA, Davis CH. 2017. Training Deep Convolutional Neural Networks for Land–Cover Classification of High-Resolution Imagery. IEEE Geoscience and Remote Sensing Letters **14**:549–553.

Schneider, C., Rasband, W. and Eliceiri, K., 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, **9**:671-675.

Singh, S.S., Singh, T.T., Devi, H.M., Sinam T. 2012. Local contrast enhancement using local standard deviation. International Journal of Computer Applications. **47**: 15.

Sladojevic S, Arsenovic M, Anderla A, Culibrk D, Stefanovic D. 2016. Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. Computational Intelligence and Neuroscience **2016**:1–11.

Tayara H, Chong K. 2018. Object Detection in Very High-Resolution Aerial Images Using One-Stage Densely Connected Feature Pyramid Network. Sensors **18**:3341.

Taylor M.E., Stone, P. Transfer learning for reinforcement learning domains: a survey. 2008. Journal of Machine Learning Research **10**:1633-1685.

Terry J.C.D., Roy, H.E., August, T.A. 2020. Thinking like a naturalist: enhancing computer vision of citizen science images by harnessing contextual data. Methods in Ecology and Evolution. **11**: 303-315.

Tzutalin. LabelImg. 2015. Git code. http://github.com/tzutalin/labelImg

U.S. Geological Survey. 2014. Early detection of invasive plants – principles and practices. Scientific Investigations Report 2012 – 5162.

Valladares F. 1999. Architecture, ecology, and evolution of plant crowns. In: F.I. Pugnaire and F. Valladares, eds. Handbook of Functional Plant Ecology. Pp. 121-194.

Vitousek PM, Walker LR, Whiteaker LD, Mueller-Dombois D, Matson PA. 1987. Biological Invasion by Myrica faya Alters Ecosystem Development in Hawai‘i. Science **238**:802–804.

Vorsino AE, Fortini LB, Amidon FA, Miller SE, Jacobi JD, Price JP, Gon Sohukani’Ōhi‘a’ōhi‘a, Koob GA. 2014. Modeling Hawaiian Ecosystem Degradation due to Invasive Plants under Current and Future Climates. PLoS ONE **9**.

Wagner, W. L., D. R. Herbst, and S. H. Sohmer. 1999. Manual of the flowering plants of Hawai`i, revised edition with 2003 supplement by Wagner, W. L. and D. R. Herbst. University of Hawai`i Press.

Wan J-Z, Zhang Z-X, Wang C-J. 2018. Identifying potential distributions of 10 invasive alien trees: implications for conservation management of protected areas. Environmental Monitoring and Assessment **190:** 739.

Wang, C.C., Samani, H. 2020. Object detection using transfer learning for underwater robot. 2020 International Conference on Advanced Robotics and Intelligent Systems (ARIS).

Wang Y, Wang C, Zhang H, Dong Y, Wei S. 2019. Automatic Ship Detection Based on RetinaNet Using Multi-Resolution Gaofen-3 Imagery. Remote Sensing **11**:531.

Weinstein, B.G. 2018. A computer vision for animal ecology. Journal of Animal Ecology. **87**: 533-545.

Weinstein BG, Marconi S, Bohlman S, Zare A, White E. 2019. Individual tree-crown detection in RGB imagery using semi-supervised deep learning neural networks.

Weinstein B.G., Marconi S, Bohlman S, Zare A, Singh A, Graves SJ, White, E. 2021. NEON Crowns: a remote sensing derived dataset of 100 million individual tree crowns. eLife: Tools and Resources.

Weiss K.R., Khoshgoftaar, T.M. 2017. Comparing transfer learning and traditional learning under domain class imbalance. 2017 International Conference on Machine Learning and Applications (ICMLA). 337-343.

West R.M. 2021. Best practice in statistics: use the Welch's t-test when testing the difference between two groups. PubMed. 58(**4**): 267-269.

Wu Z, Chen Y, Zhao B, Kang X, Ding Y. 2021. Review of weed detection methods based on computer vision. Sensors. **21**: 3647.

Yosinski J, Clune J, Bengio Y, and Lipson H. 2014. How transferable are features in deep neural networks? Advances in Neural Information Processing Systems 27 (NIPS '14): 1-14.

Yuan W, Wijewardane NK, Jenkins S, Bai G, Ge Y, Graef GL. 2019. Early Prediction of Soybean Traits through Color and Texture Features of Canopy RGB Imagery. Scientific Reports. **9**:1-17.

Zhao ZQ, Zheng P, Xu ST, Xindong W. 2019. Object Detection With Deep Learning: A Review. IEEE Transactions on Neural Networks and Learning Systems. **99**: 1-21.

Zou Z, Shi Z, Guo Y, Ye J. 2019. Object detection in 20 years: a survey. Computer Vision and Pattern Recognition.