APPLYING SEGMENTATION AND NEURAL NETWORKS TO DETECT AND QUANTIFY MARINE DEBRIS FROM AERIAL IMAGES CAPTURED BY AN UNMANNED AERIAL SYSTEM AND MOBILE DEVICE

A Thesis

by

KIEU TRAN

BS, Unity College, 2014

Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in

COASTAL AND MARINE SYSTEM SCIENCE

Texas A&M University-Corpus Christi
Corpus Christi, Texas

August 2018

August 2018

APPLYING SEGMENTATION AND NEURAL NETWORKS TO DETECT AND QUANTIFY MARINE DEBRIS FROM AERIAL IMAGES CAPTURED BY AN UNMANNED AERIAL SYSTEM AND MOBILE DEVICE

A Thesis

by

KIEU TRAN

This thesis meets the standards for scope and quality of
Texas A&M University-Corpus Christi and is hereby approved.

Jeremy Conkle, PhD
Chair

Michael Starek, PhD
Committee Member

James Gibeaut, PhD
Committee Member

August 2018

ABSTRACT

Marine debris is a global issue with adverse impacts on the marine environment, wildlife, economy, and human health. Its presence on beaches may vary due to topography, hydrological conditions, proximity to litter sources, and the extent of beach use. Studies of these parameters on beach litter are essential for understanding spatial and temporal patterns; however, this work is labor-intensive and time-consuming. To overcome these obstacles while gaining higher spatial and temporal resolution data, two methods were developed (1) segmentation and (2) regions with convolutional neural networks (R-CNN) to detect and quantify marine macro-debris using high-resolution imagery.

Data to develop and test the methods were collected using a rotary Unmanned Aerial System (UAS) with an RGB sensor at various altitudes over a 100m section of sandy beach on Mustang Island, Texas. Images were processed through structure-from-motion photogrammetry to derive orthomosaics for each flight. Orthomosaics were then run through an image processing, and classification workflow developed for segmentation and delineation of imaged debris. The segmentation algorithm detected the most debris at the lowest altitude (215 of 341 total at 15m), with decreasing detections at 22m (101) and 35m (50). The second method applied deep learning object detection to smartphone images. Based on the dataset and the network architecture, R-CNN mean average precision can range from 31.4% to 66% (Girshick et al., 2014). Training for R-CNN consisted of three stages: extract region proposals, train AlexNet to classify objects, and train a bounding box regression model to locate the debris. Average precision for the Specifically Engineered Algorithm for Gathering and Understanding Litter Location (SEAGULL) detector was ~22%, meaning it has a low performance at detecting all the debris in the testing dataset and

correctly predicting whether or not that region was debris or not. Object detection has been a challenging task for decades, and with a low overall debris detection accuracy, it needs to be further improved with the use of a larger dataset or adjusting the training parameters. Few studies have been published on this topic, but this work demonstrates that remote sensing with UAS has the potential to increase research efficiency.

DEDICATION

To Mrs. Mills, Mr. Mac, and Dr. Ongley; my teachers. I have met a lot of brilliant teachers while I was working towards my Master's degree, but these three individuals helped shape me into the person I am now and have inspired me to pursue two degrees in coastal and marine science.

Mrs. Mills, without you, there was no way I could have been a Teaching Assistant for Chemistry I & II and Organic Chemistry II. Thank you for putting up with us and teaching us in ways that are engaging and fun. I know you have inspired many students to challenge themselves and work hard towards their goals and dreams, including me. I am incredibly thankful that our high school had such a great teacher; one that does not give up on her students.

To Mr. Mac; my teacher, mentor, therapist, and life coach. Thank you for all the after school conversations and listening to me. You were the one that pushed me out of the darkness, and you were the one that inspired me to pursue my B.S degree in Maine. Without you, I would not have stepped out of my comfort zone and explore what the world has to offer.

Finally, to Dr. Ongley, my first advisor. In our first meeting, you mentioned the Research Experience for Undergraduates (REU) internships and how competitive it was. I thought I was not qualified for it because I was only a freshman at that time and I did not have enough experience. Because of your help and recommendation, I got my first REU. From then on, I was not afraid of rejections or competing with someone with more experience. I may be the only women or Asian American in the group, but I am just as qualified as anyone else and you taught me that.

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor and graduate committee chairman, Dr. Jeremy Conkle. Without your guidance, support, and advice throughout this process, I would not have a thesis or have learned so much. Many thanks to my committee members, Dr. Michael Starek and Dr. James Gibeaut for agreeing to serve on my committee and providing helpful feedback and the resources I need for this research. I would also like to thank the MANTIS Lab and Jake Berryhill. Thank you, Jake, for coming out in the field with me and teaching me so much about UAS and Pix4D. To the Texas Comprehensive Research, thanks for the funds which supported this research. Thank you to my friends, colleagues, and TAMUCC staff for making my time at Texas A&M University-Corpus Christi an enjoyable experience and less stressful.

Finally, I want to thank my husband, William Nguyen. Thank you for pushing me to pursue my dreams and supporting me. Without you, I would have been stuck in Florida and never left. You are my “pirate king” (One Piece reference).

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES PAGE

LIST OF TABLES

TABLES PAGE

LIST OF ABBREVIATIONS

R-CNN	Region with Convolutional Neural Network
CNN	Convolutional Neural Network
UAS	Unmanned Aerial System
FAA	Federal Aviation Administration
SfM	Structure-from-Motion
DSM	Digital Surface Models
Greedy NMS	Greedy Non-Maximum Suppression
SVM	Support Vector Machine
SEAGULL	Specifically Engineered Algorithm for Gathering and Understanding Litter Location
NIR	Near-Infrared
LIDAR	Light Detection and Ranging
YOLO	You Only Look Once

CHAPTER I: INTRODUCTION

1.1 Marine Debris

Contamination of marine debris is a global problem, and increasing sea level, wind speed, and oceanic currents associated with climate change are likely to influence the transport and fate of this debris on beaches, in estuaries, and in the open ocean (Browne et al., 2015). According to the Marine Debris Research, Prevention, and Reduction Act, "marine debris is defined as any persistent solid material that is manufactured or processed and directly or indirectly, intentionally or unintentionally, disposed of or abandoned into the marine environment or the Great Lakes" (Marine Debris Research, Prevention, and Reduction Act, 33 USC §§ 1951). Our ocean and coastal systems are polluted with a wide variety of marine debris; common items are cigarette butts, plastic bottles, plastic caps, and food wrappers (Belhouari et al., 2017). Marine debris is items that have been made or used by people and is either from inland sources or transported to the coast by rivers, wind, human-made drainage systems or human activity. (Corcoran et al., 2009; Duhec et al., 2015). It has adverse impacts on the marine environment, wildlife, economy, and human health (Engler, 2012; Critchell et al., 2015). Marine debris has negatively impacted marine animals of all trophic levels, through entanglement, ingestion, habitat destruction, and physiological effects (Duhec et al., 2015; Gregory, 2009). Debris can also modify coastal habitats and transport invasive species. It can also have negative economic impacts, which include damage to vessels when floating (nets, solid objects, and even plastic bags) or submerged and the loss of tourism revenues due to a reduction in the visual appeal of a destination (Jang et al., 2014). Human health may also be affected, either by physical injury, transfer of toxic chemicals into the marine food web and to human diet, and ingesting synthetic debris from tap water, beer, and sea salt (Ribic et al., 1992;

Browne et al., 2007; Yang et al., 2011; Engler, 2012; Kosuth et al., 2018). Consequently, there is a need to address the growing presence of marine debris.

The concentration of marine debris varies among locations, which include both densely populated regions to remote areas. The amount of marine debris at each site is influenced by a number of variables, such as the physical characteristics of the debris, hydrology, human activities, and oceanographic processes (Browne et al., 2015). These variables can affect the transportation process, transport direction, and accumulation rates of marine debris. Marine debris is sometimes distinguished by its source, either land-based, sea-based or both (Thiel et al., 2013). Land-based sources originate from human activities, such as recreational use of beaches, landfills, sewage, and debris that are washed into coastal waters by natural disasters such as hurricanes and tsunamis. Sea-based sources include shipping, fishing, and oil and gas extraction activities. However, it can be challenging to identify the source or the origin of the debris, especially when the floating debris sinks, fragments into smaller pieces, and is consumed by organisms or deposited in the sediment (Thiel et al., 2013). Marine debris is classified based on size and type (Ribic et al., 1992). According to NOAA Marine Debris Monitoring and Assessment, macro-debris is greater than 2.5cm, meso-debris is from 5mm -2.5cm, and micro-debris is less than or equal to 5mm (Lippiatt et al., 2013). NOAA also classified the debris based on material type, which includes plastic, metal, glass, rubber, processed lumber, cloth/fabric, and other/unclassifiable material (Opfer et al., 2012).

Plastic is the dominant type of marine debris, accounting for 60-80% of material found on beaches, (Derraik, 2002; Lee et al., 2013; Browne et al., 2015). Annually, ~4.8-12.7 million metric tons of plastic debris enter the ocean from coastal zones or ships and has significantly increased over the last five decades (Jambeck et al., 2015). In the environment, plastic debris can fragment into smaller particles, which can increase their spatial distribution (Lee et al., 2013). Beach

topography, near-shore currents, storm activity, proximity to litter sources, and visitation rates all affect the amount and type of debris that is present (Corcoran et al., 2009). Lightweight, buoyant materials such as plastics bottles and utensils that float are circulated by ocean winds and surface currents until they sink into the water column or settle on a coastline.

In the Gulf of Mexico, the major current is called the Loop Current. It is an area of warm water that travels northward from the Caribbean at a speed of 0.8 m/s, reaching the northern Gulf of Mexico before looping back south and flowing into the Atlantic Ocean through the Florida Straits. Eddies periodically spin-off from the Loop Current and travel westward towards the Texas coast at a speed of about 2-5 km/day (Xu et al., 2013). The Loop Current captures marine debris along its path, and its eddies transfer some of this material to the western continental shelf of the Gulf of Mexico. The material is then captured in longshore transport cells that run parallel to the shore, helping to deposit marine debris on coastal beaches and marshes (Figure 1). Padre Island National Seashore can accumulate up to one ton of marine debris per linear mile, with approximately 90% being plastics (National Park Service, 2005). This buildup of debris on Padre Island National Seashore varies during the year due to the wind and current patterns of the Gulf, weather conditions, and sea-based activities (Ribic et al., 2011).

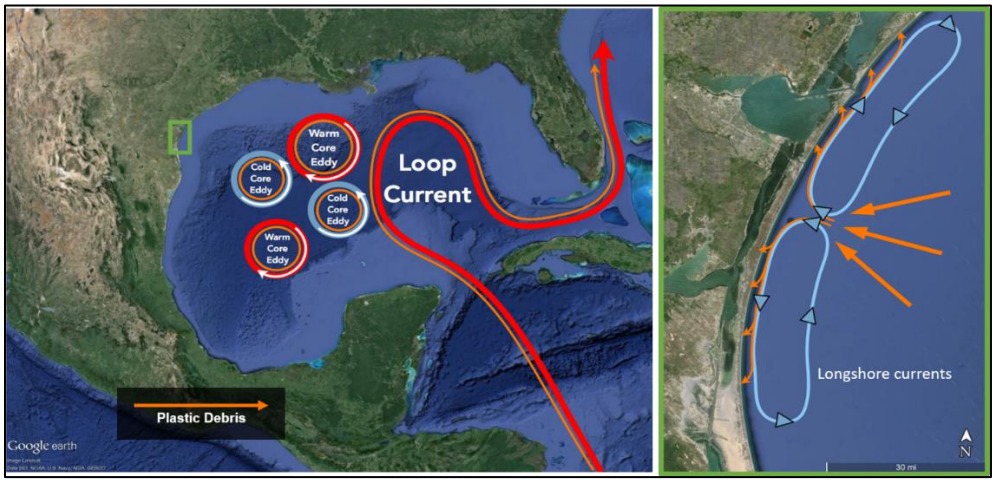


Figure 1: Concentration of marine debris varies due to the influence of the Gulf Loop current, eddies, and longshore currents (Illustration done by Dr. Conkle using Google Earth).

1.2 Beach Surveys

With the high accumulation rates of marine debris on beaches, like Padre Island National Seashore, characterizing and cataloging this material as well as studying its presence and movement is labor-intensive and time-consuming. In the last decade, there has been an increase in published papers involving marine debris and different approaches to gathering marine debris data such as sources, composition, amounts, and environmental significance based on monthly or longer in situ beach surveys (Browne et al., 2015; Galgani et al., 2015). Beach surveys using conventional surveying techniques are the most widely used method for monitoring the amount and type of litter in a specified area at a specific time or how litter changes with time (Rubic et al., 1992). Various groups like NOAA Marine Debris Program and Ocean Conservancy have developed a standardized marine debris shoreline survey protocol to allow trained individual and volunteers to gather valuable data for monitoring and assessing marine debris (Opfer et al., 2012). The traditional beach surveys rely on visual methods, where surveyors count and classify individual debris items, the debris may or may not be removed from the beaches, and either the entire beach or smaller sections (100m transects) may be surveyed (Ribic et al., 1992). In addition to the high expense, labor and time required they also cover varying spatial and temporal scales making comparisons across studies difficult (Jang et al., 2014). To overcome the obstacles of conventional monitoring, researchers used webcams and applied image analysis techniques to sequentially quantify debris abundance at specific sites (Kako et al., 2010; Kataoka et al., 2012). A remote-controlled digital camera suspended from a balloon was used to estimate the total mass of litter washed ashore on beaches (Nakashima et al., 2011). An unmanned aerial system (UAS) model GAUI 500X was used to develop a beach litter detection software for debris larger than 50cm (Jang et al., 2014; Jang et al., 2015). Light Detection and Ranging (LIDAR) was used to classify and quantify marine debris

based on its shape (Ge et al., 2016). Moreover, a single-engine, general aviation aircraft with fixed landing gear was used to manually detect macro-debris ranging from $<0.5m^2$ to $>2.0m^2$ (Moy et al., 2017). These approaches, while potentially advantageous over traditional methods, do not provide as much promise as a small, low-flying, inexpensive Unmanned Aerial Systems to study marine debris larger than 2.5cm on beaches.

1.3 Small Unmanned Aerial System (UAS)

Small UAS or drones are remotely or self-piloted aircraft weighing less than ~25kg with a payload that includes cameras and various sensors. This technology is rapidly advancing due to a growing demand for their use in military, civil, hobby, and research (Itkin et al., 2016). For research, UAS is used to collect data in a variety of fields like agriculture, animal behavior and population size, and coastal mapping. Applications of using UAS for coastal and marine surveys include dune and beach face erosion measurement, detecting sizeable marine debris at sea, and seal population surveys (Turner et al., 2016; Veenstra & Churnside, 2012; Seymour et al., 2017). Vertical aerial photographs are the most commonly used data source in shoreline mapping since they provide a valuable record of shoreline position. However, distortion and displacement are greater in small-scale aerial photographs (e.g., 1:50,000); those imageries cover large areas with less detail. Low-flying UAS can produce a higher spatial resolution at a lower cost, a closer view of the areas of interest, and used in a variety of situations (Ramesh et al., 2016). Due to the low flying ability of a UAS and the high-resolution imagery, detecting macro-debris is possible. There are two main types of UAS's used for mapping purpose: fixed wing and rotary wing; each with its advantages and disadvantages. Despite being limited by meteorological conditions (wind speed and rain), battery life, and the flight restriction areas, UAS is useful for coastal research and monitoring. The flexibility of UAS will reduce cost at local geographic scales (e.g., routine monitoring a few km

stretch of beach) and offer better spatial and temporal resolution than conventional field survey approaches to marine debris monitoring. Manually counting marine debris in imagery is time consume and inefficient. However, recent advances in automated counting and computer vision approaches can help overcome these inefficiencies.

1.4 Object Detection

Object detection is a growing field of study related to computer vision, machine learning, and digital image processing. A good detector can identify the object, categorize it and provide the location of the object (Chen et al., 2018). For humans, we can effectively recognize and localize objects in a scene or image, but for artificial systems, recognition is challenging due to factors, such as shape, distortion, texture, and color (Dan et al., 2014). Different techniques for solving this problem have been proposed, ranging from low-level to high-level processing.

Digital image processing performs various operations that improve image quality and extracts relevant information (Chudasama et al., 2015). Low-level processing improves the quality of the image through enhancement or noise removal processes. Mid-level processing can extract some attributes from an image through segmentation, feature-based, and object recognition. Image segmentation tries to segment or separate the part of an image that contains the target objects from the image background or other targets. High-level processing understands and recognizes objects in an image for analysis; such as using support vector machine classification (SVM) with a histogram of oriented gradient features, convolutional neural network (CNN), and other computer vision techniques (Gonzalez and Woods, 2007).

High-level processing is advancing rapidly due to “deep learning” that has been widely applied to various computer vision tasks such as image classification, object detection, and scene labeling (Sun et al., 2017). Deep learning is a machine learning technique that enables the computer

to build and understand complex concepts out of simpler concepts. Deep learning is usually implemented using a neural network architecture, such as CNN. CNN typically consists of multiple layers including the convolutional layers and each layer has many neurons with learnable weights and biases. With each layer, the network will start to recognize patterns, such as specific color, edges, shape, and texture. To achieve high predictive accuracy from a CNN, it must be trained with large datasets; instead, this project used transfer learning with an existing pre-trained network (AlexNet) and modified it with a project-specific dataset (Pan & Yang, 2010). Transfer learning has been used in a variety of deep learning application, and it is time efficient than training the network from scratch. The modified network was then further re-trained using Region-Based with CNN (R-CNN) to classify and locate the objects in the image.

1.5 Objectives

The primary purpose of this research is to increase the efficiency and temporal repeatability of marine debris monitoring relative to standardized field counting approaches through remote sensing techniques. To achieve this goal, the first objective for this project is to assess the impact of flight design and ground sample distance on detection and quantification of marine debris. The second objective is to design and established the parameters for segmentation by color thresholding and R-CNN to make a plausible estimate of the quantities and locations of macro-debris deposited on sandy beaches from aerial images acquired from a small rotary UAS and smartphone.

CHAPTER II: METHODS AND MATERIALS

2.1 Study Area

Federal Aviation Administration (FAA) guidelines limited the potential beach survey study sites in the local area due to the Corpus Christi Naval Air Station, Padre Island National Seashore, and airports/heliports (Figure 2). These rules state that operating a UAS for research is considered recreational use. However, the pilot must follow FAA guidelines to minimize risks to other aircraft, people, and property. These include aircraft weighing less than ~25kg, flying at or below ~122m, keeping the UAS within sight, never flying near other aircraft and airports, and never flying over unprotected people on the ground who are not directly participating in the UAS operation (Recreational Users, 2015).

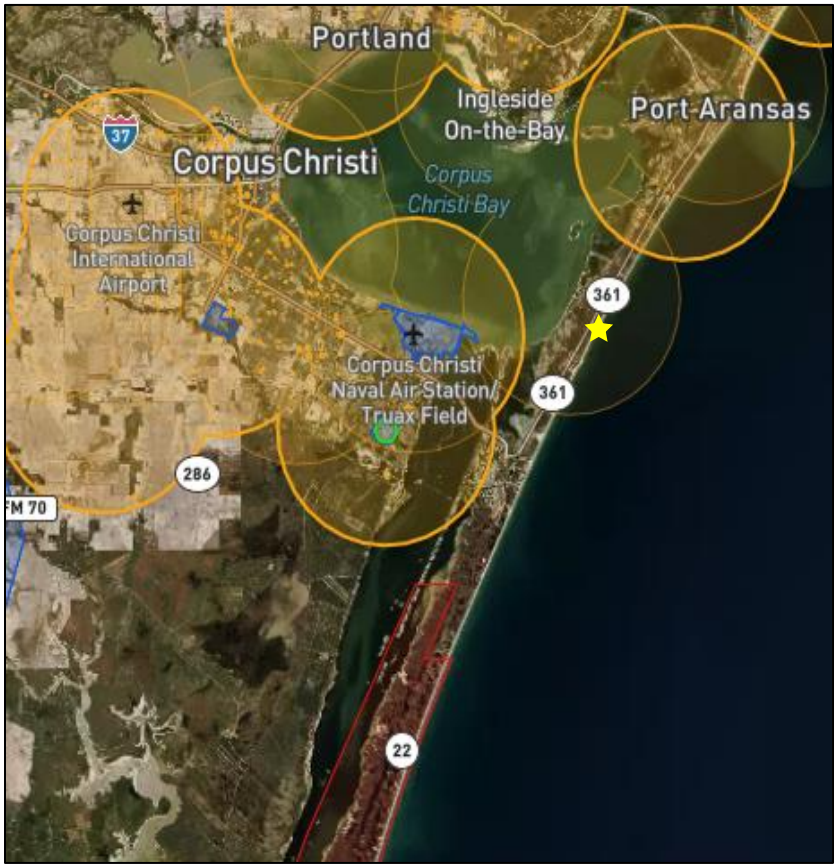


Figure 2: Restricted areas (red), special use airspace (blue) and advisories (orange) of where to fly a UAS around Corpus Christi, TX (Picture from AirMap). The study site (yellow star) location is in an advisory area.

The area chosen for this research was a 100m section of beach on Mustang Island, TX (Figure 3). The specific study area was chosen to meet FAA safety guidelines, located away from any access points, and minimum influence from human activities. However, Corpus Christi Naval Air Station had flight restrictions due to security reasons, and flight plans were canceled or rescheduled more than once. To obtain the best imagery possible, the flights were scheduled once a month around low tide to capture the majority of the beach shoreline, and it was typically taken early in the day to avoid foot traffic and parked vehicles on the beach. Flights were also limited by wind speed; the preferred wind speed was less than 17mph to avoid turbulence, improve handling of the UAS, ensure public safety, and reduce battery usage per flight.

Figure 3: 100m section of beach (yellow star) on Mustang Island, TX. The focus area starts from the dune toe to the low tide shoreline (Map from Google Earth).

2.2 UAS

Aerial images were collected using two different UAS. First was the DJI Phantom 4 Pro (Figure 4A) with a 1-inch 20-megapixel CMOS sensor, weighing ~1.4kg including the battery and propellers, and maximum flight time of ~30 minutes (Phantom 4 Pro, n.d.). When the Phantom 4 Pro was not available, the 3DR Solo with a Sony QX camera with similar properties was used (Figure 4B). A rotary UAS was used rather than a fixed-wing UAS because of its capability of

holding a steady field of view for an extended period, no need for a launching and landing runaway, and ability to handle highly dynamic areas, like the beach.

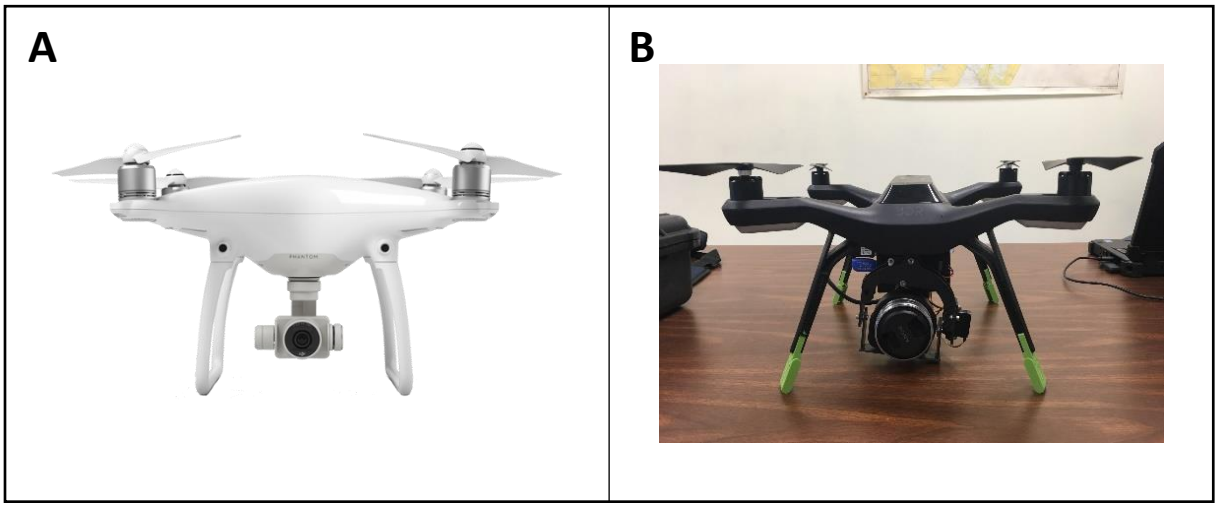


Figure 4: The DJI Phantom 4 Pro (A) was used for the majority of the surveys, and the 3DR Solo (B) was a backup.

2.2.1 Mission Plan

The mission plan (flight path, height, speed, image overlap and side-lap) varied for each flight. Plans were developed with Ground Station Pro (GS Pro), an iPad app designed to control and plan automatic flights for the DJI Phantom 4 Pro. This app also provides real-time flight information such as location, remaining battery life, and progress. The 3DR Solo used Mission Planner, an open source software similar to the GS Pro.

Choosing a flight altitude is critical for debris visibility, debris detection, and finding a balance between image resolution and coverage. To determine optimal flight altitude, ground control targets were used as a standardized measurement, and because macro-debris is considered as items larger than 2.5cm (Lippiatt et al., 2013); the flight altitudes were chosen based on the clarity of the 7.6cm diameter circle of the target (Figure 5). For this UAS application, optimal flight altitudes were between 15 and 35m (Figure 6). Flights ≥20m were completed with one

battery and in one mission plan, while flights ≤20m required multiple batteries usage and mission plan to cover the area.

Each mission plan was implemented with a high overlap ratio (≥70%) to enable structure-from-motion (SfM) photogrammetry capability. Flight plans were uploaded to the UAS via Wi-Fi connection between the remote controller and the UAS. During the autonomous aerial survey, all images captured were geo-tagged by the onboard GPS as required for SfM processing. All the aerial images and data were stored in the platform's onboard SD card and later transferred to a computer for storage and processing.

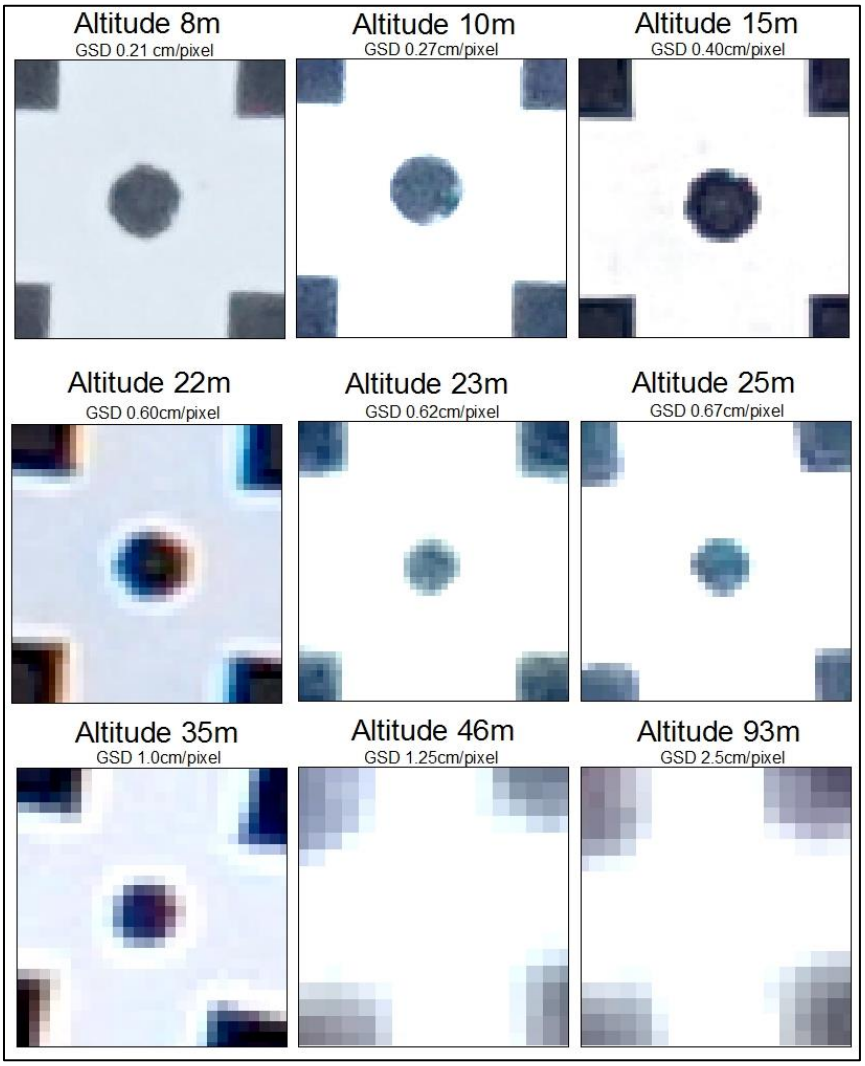


Figure 5: Target center (7.6cm) was compared across different flight altitudes and ground resolutions to determine the best flight plans.

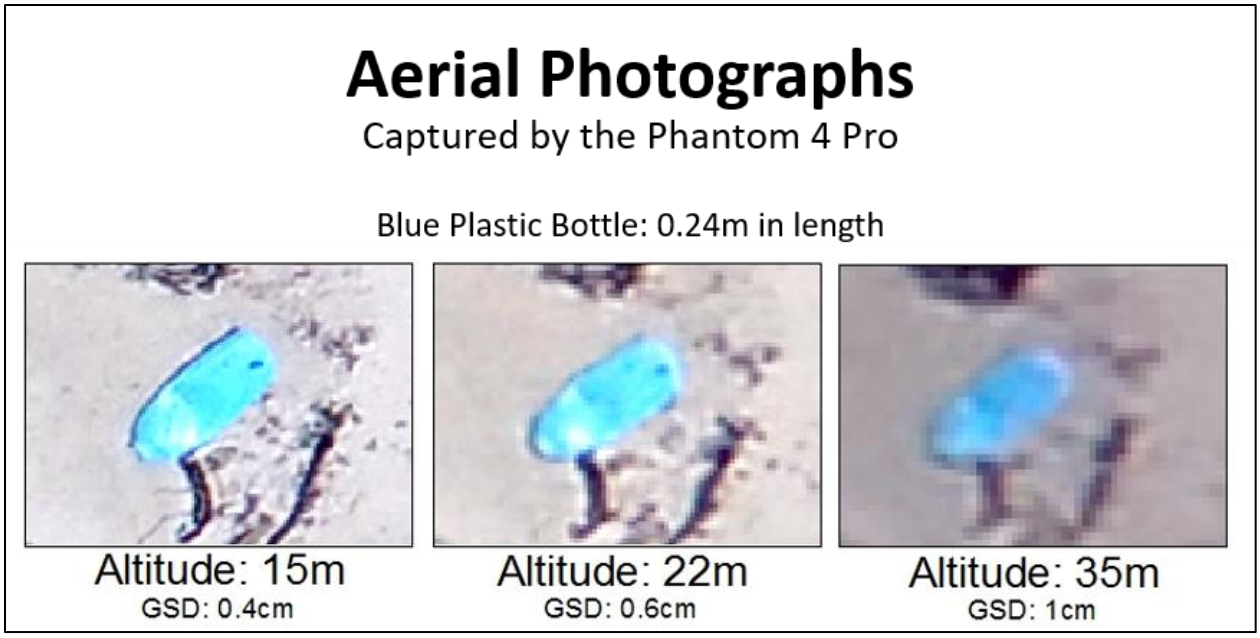


Figure 6: Aerial photographs of a blue plastic bottle comparing the ground sample distance and flight altitude.

2.2.2 Ground Control Targets

Ground targets were used to georeferenced images for time series analysis at the site. This simplifies image alignment with the chosen coordinate system (WGS84 UTM Zone 14N), and reduces shifting, rotating, or skewing of images. Each ground control target was surveyed with an Altus APS-3 Real Time Kinematic GPS with a Leica CS10 field controller. WGS84 UTM Zone 14N was the horizontal coordinate system used for this survey with heights (z-values) referenced to North American Vertical Datum 1988 computed by transforming WGS84 ellipsoid heights using the National Geodetic Survey GEOID 12b model. The northing and easting coordinates, as well as the elevation for each target, were exported to a comma separated value text file and saved for use during image processing. This project used 5, 2.5 x 4ft black plywood boards painted with a white cross, and a black circle (7.6cm diameter) in the center (Figure 7). Four control targets were set in the corner of the beach and one in the center, creating a symmetrical pattern as to maximize the geolocation accuracy across the entire beach section (Figure 8).

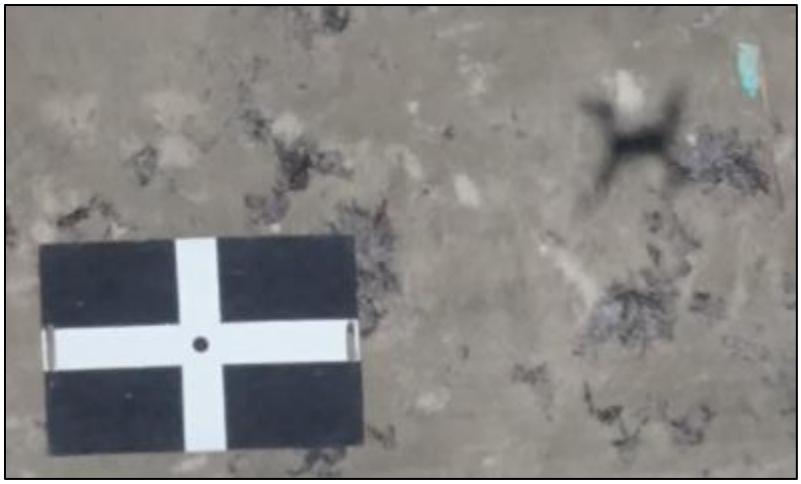


Figure 7: A ground control target used during the beach survey.

Figure 8: Location of the ground control targets (orange circle) used across a 100 m section of beach.

2.2.3 Ground Truth Data

To compare UAS detection of plastic materials with actual material on the beach, all macro-debris (larger than 2.5cm) was geotagged (Figure 9) using a Garmin Monterra handheld GPS and then imported into ArcMap to create a point shapefile. Since the Garmin Monterra has a GPS/GLONASS receiver, the GPS positioning of the debris were checked using the raw images and orthomosaic. These data were also used later in the project to compare the number of debris detected through segmentation, which divides the image into multiple parts (foreground and background) to help identify marine debris.

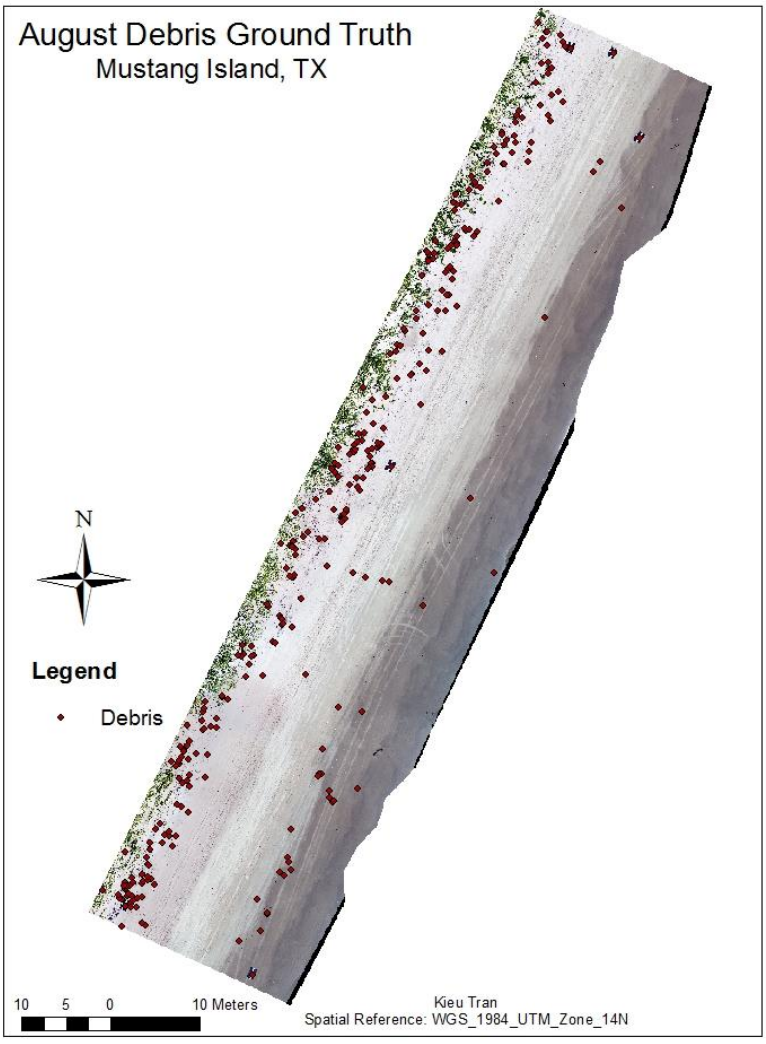


Figure 9: Ground truth map of debris (341 items; red diamonds) on a 100m section of Mustang Island, TX beach.

2.3 Data Processing

SfM is a computer vision technique that relies on a series of overlapping, offset images to recreate an environment (Westoby et al., 2012). The georeferenced SfM datasets consist of the X, Y, and Z (latitude, longitude, and elevation) information with additional RGB spectral (red, green, and blue) (Mathews & Jensen, 2013). Post-flight, images were processed using Pix4D SfM software, to create an orthomosaic for each data collection. Pix4D is commercial software that implements SfM photogrammetry to convert the aerial images into georeferenced 2D or 3D surface models and point clouds. It was chosen for this project due to it simple-to-use interface and ability to process images taken from varying angles and distances to create 2D maps and 3D models.

2.3.1 Pix4D Settings

Images were loaded into Pix4D and the appropriate input and output coordinate systems, as well as geolocation accuracy, were determined. The output coordinate chosen for this project was "Geoid Height above WGS84 Ellipsoid" to match the ground control point coordinate system. Pix4D has three processing stages: (1) initial processing, (2) point cloud and mesh, and (3) Digital Surface Models (DSM), orthomosaic, and index. Stage 1 begins matching image features and generates a sparse point cloud. The processing options were set to default except for the Matching, which was set to "Triangulation of Image Geolocation." This was chosen because the aerial images were geo-tagged and the images were matched with at least three similar images. The Matching Strategy was selected for this project, using the "Geometrically Verified Matching" to look at relative camera positions and discard geometrically unrealistic matches. After this processing stage, Pix4D generates a "Quality Report" and a low-resolution orthomosaic preview. This report summarized the project and provided a quality check to the images, dataset, camera optimization, matching, and georeferencing before starting stage 2.

After stage 1, a ground control coordinate text file was imported into the project with the specified coordinate system of the control points. Using the RayCloud Editor in Pix4D, two images of each target were manually tagged, which then enable the software to tag the remaining images containing the ground control targets automatically. All images containing a ground target were manually inspected to make sure they were centered before they were rematched and optimized. The addition of ground control points improved the quality and accuracy of the orthomosaic.

For stage 2, Point Cloud and Mesh increase the density of 3D points in the point cloud and 3D textured mesh. Users can update image processing options and desired outputs for the generated point cloud. In this stage, the parameters were set to default, producing denser 3D points for the point cloud to render a more accurate DSM for creating the orthomosaic.

Lastly Stage 3, DSM, Orthomosaic, and Index, the users finalize the processing options and outputs for the DSM and orthomosaic. The resolution, DSM filter, raster DSM, and orthomosaic was set to default to reduce noise and to improve the image quality by removing redundant points. The orthomosaic was saved as a single orthomosaic GeoTIFF file without transparency for digital image processing.

2.4 Mid-Level Processing

After initial image processing, mid-level and high-level processes were performed to detect marine debris captured in the orthomosaics. Matlab (R2017b) was used to perform the remaining image processing (segmentation, morphological operation, and extraction) and apply deep learning (transfer learning, R-CNN, and evaluation) to the images. Due to its large file size, the orthomosaic file was evenly split into four tiles before processing (Figure 10). In Matlab, the RGB images are stored as an *m*-by-*n* by 3 data array that defines red, green, and blue color components for each pixel. The *m*-by-*n* grid of pixels represents *m* as the number of columns (width) and *n* as the

number of rows (height) in an image. For the mid-level processing, each tile went through three stages: Image Preprocessing, Morphological Operations, and Extract Features (Figure 11). The parameters applied during these three stages were manually determined through trial-and-error.

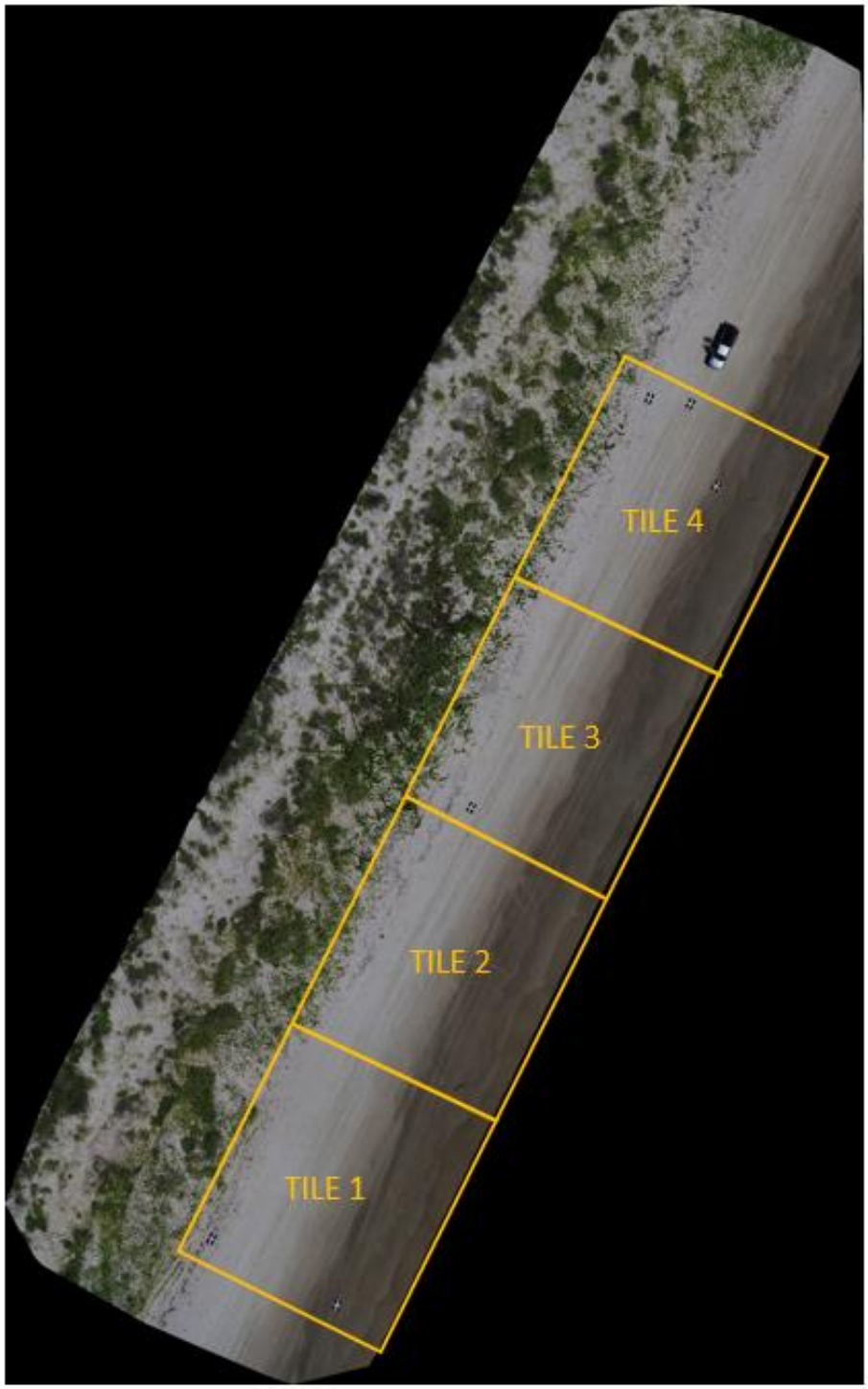


Figure 10: Site was divided into four even tiles for Matlab to process.

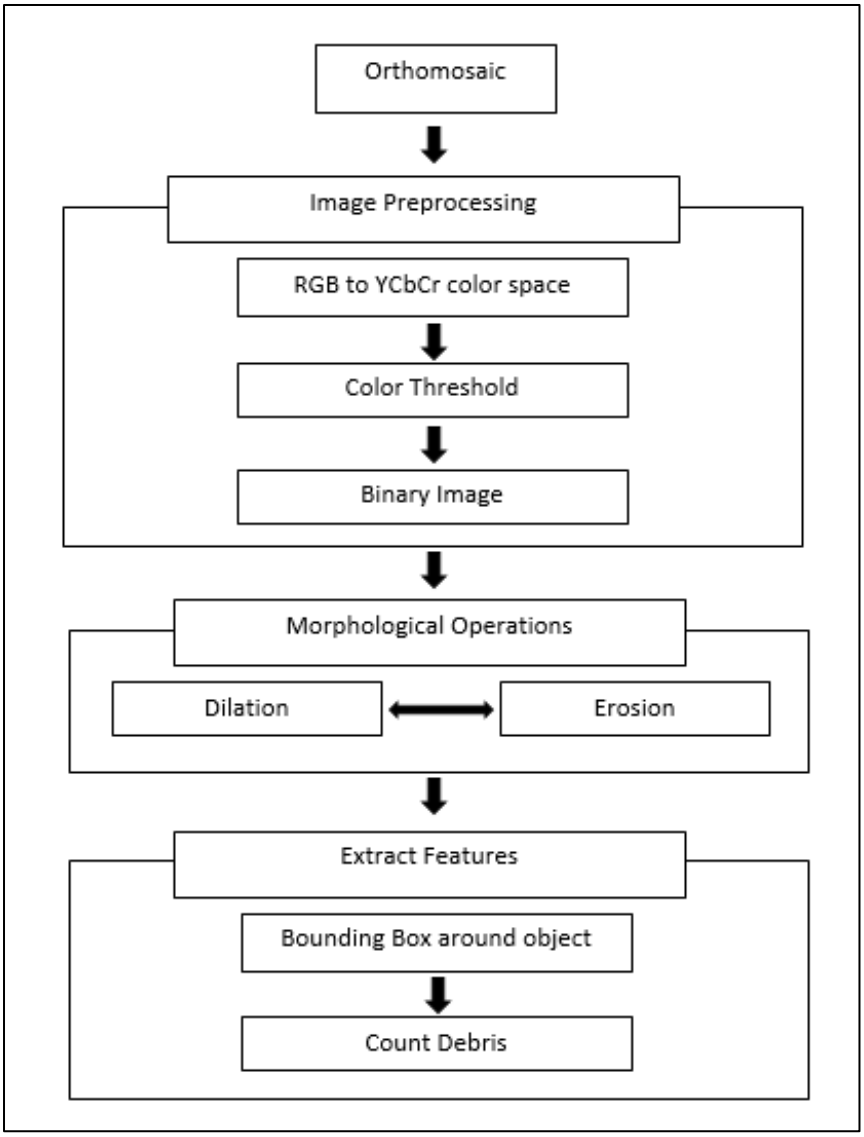


Figure 11: Method flowchart for detecting and counting marine debris using segmentation color thresholding.

Image Preprocessing involves using Matlab Color Thresholder Application, an interactive application that automates colormap conversion, color threshold, and binary image conversion. This was done to convert the RGB image to YCbCr. RGB color space defines the color as the percentage of the mixture between red, green, and blue hues; a standard color space used in image processing. However, other color spaces describe colors by their hue, saturation, luminance, or intensity (Figure 12). By testing out each color space, the best approach to detect the debris was through YCbCr.

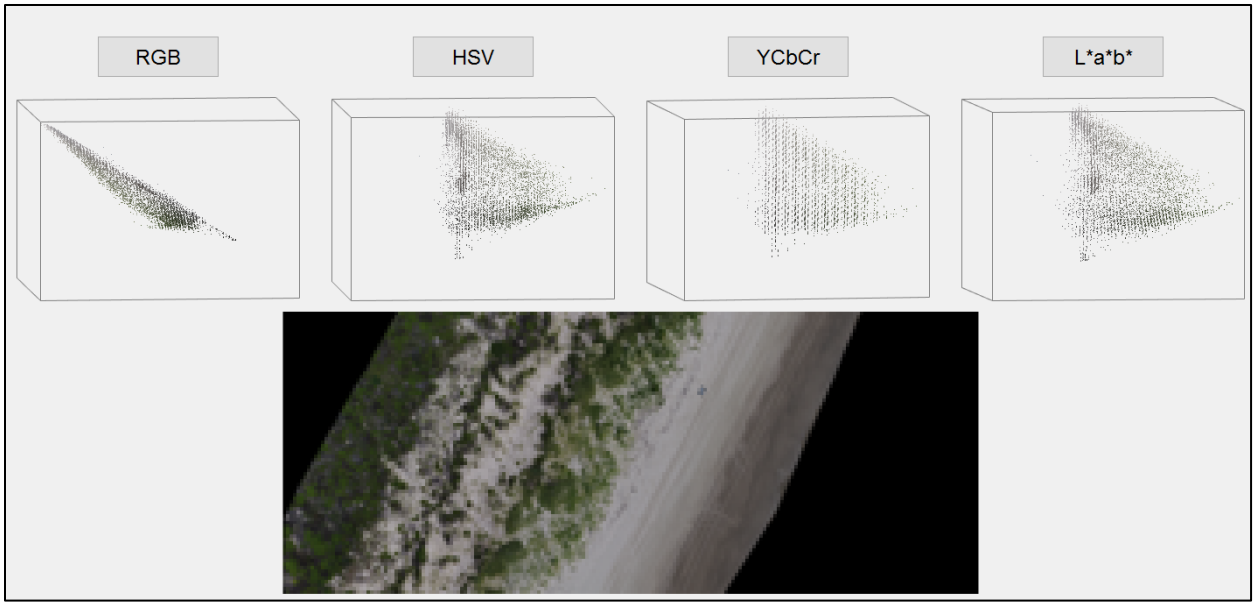


Figure 12: Color Thresholder Application of Matlab is displaying the various color space for this aerial image.

With YCbCr, Y, Cb, and Cr represent the brightness, the blue-yellow spectrum, and the red-green spectrum. YCbCr was also chosen because it could exclude the background pixels that have no value ([0 0 0]) without affecting the orthomosaic. To automatically detect the pixels containing the targeted debris in the images, the color space threshold was defined (Table 1). Segmentation by color thresholding used the color information of the object to extract it from the background and other objects. This technique specifies the range of YCbCr intensities for thresholding. If the pixel values were outside the range for all three channels, that pixel would be considered as the background (sand, vegetation, and water) and pixels that are within the threshold would be considered as the foreground (debris) (Table 1). The color threshold for each tile varies due to the amount and angle of the sunlight, but they are all close in range. Using the "select a region of colors in the image" tool in Matlab, the majority of the background pixels were excluded, and three thresholds were established (Figure 13). The image was exported as a binary image (Figure 14A), a masked image, and the function was exported and saved.

Table 1: Color threshold range for each altitude based on histogram settings.

Altitude	Tile Number	Channel 1 Threshold (min/max)		Channel 2 Threshold (min/max)		Channel 3 Threshold (min/max)	
35	1	16	215	89	133	116	139
35	2	16	197	93	133	118	149
35	3	16	195	67	135	116	154
35	4	16	204	77	136	113	154
22	1	16	184	85	134	116	140
22	2	16	217	93	136	115	140
22	3	16	211	71	134	116	148
22	4	16	199	77	134	117	148
15	1	16	198	78	138	112	149
15	2	16	206	72	151	101	154
15	3	16	212	62	147	110	158
15	4	16	205	71	144	105	154

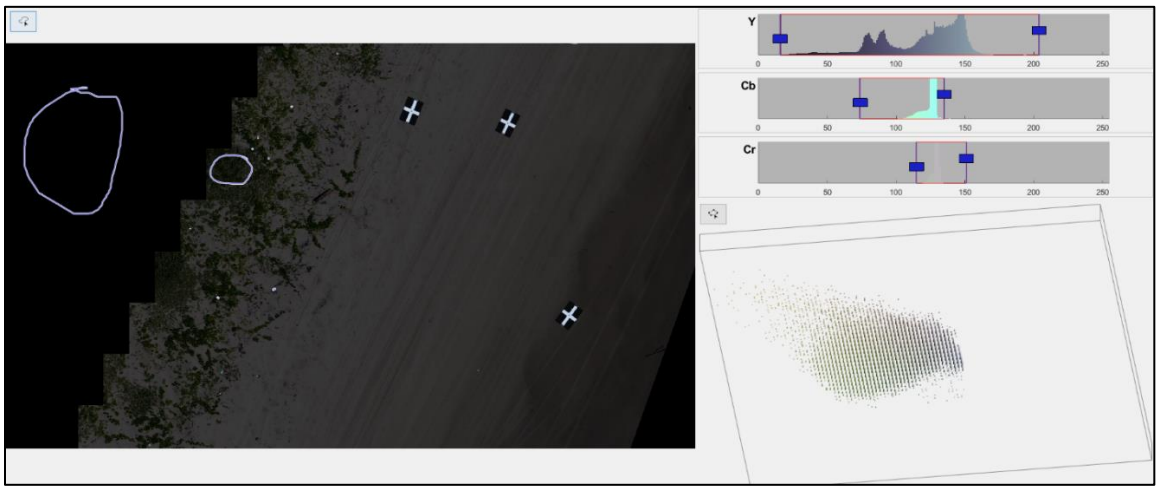


Figure 13: Removing most of the background pixels and establishing the YCbCr range.

Next is the Morphological Operation stage, which aims to eliminate the noise with minimal image distortion. The two operations used for this project were Dilation and Erosion. Dilation "expands" the boundary by adding pixels to the boundaries of objects in an image while Erosion "contracts" or removes the pixels (Figure 14B). Erosion was used first to remove the noise such as vegetation or sand pixels. After those pixels were removed, the dilation operation was used to restore the debris shape that was affected by erosion. The number of pixels added or removed from the object in an image depends on the size and shape of a predetermined structuring element. The structuring element used for this project was either "disk" or "diamond" shape, with size varying based on noise and debris present. The noise is random background pixels that are left after the majority of the background have been removed from the Image Preprocessing Stage. The structuring element shape and size for each tile was decided on whether it can remove the majority, if not all, of the noise and whether it affected the debris pixels. The center pixel of the structuring element is called the origin, and the size of the element is either the radius of the disk or the distance from the origin to the points of the diamond. The size of the structuring element used in each tile was determined through trial and error. The size started off small (1-3 pixel) before it was increased to the 5-6 pixel so that the background pixels can be removed. By increasing the structuring element of erosion, the original shape of the debris was affected. However, as long as there are enough debris pixels present, it can be dilated to restore its shape. Sometimes, the objects had holes from the Erosion stage, so an additional fill tool was used to help fill in the missing pixels (Figure 14C).

Last is the Extract Feature stage, where the debris will be counted and detected using bounding boxes (Figure 14D). To detect the debris, "regionprops" function in Matlab was used to measure properties of the binary image. Properties like centroid and eccentricity were used to

determine the center of mass of the region and to count the objects. To find the location of objects in the image, a bounding box was used to determine the [x y width height] of the object, x and y are the coordinates to the center of the object, and the width height is the size of the box. The detection rate was calculated for each altitude by measuring the proportion of true positive to the sum of true positive and false negative detection.

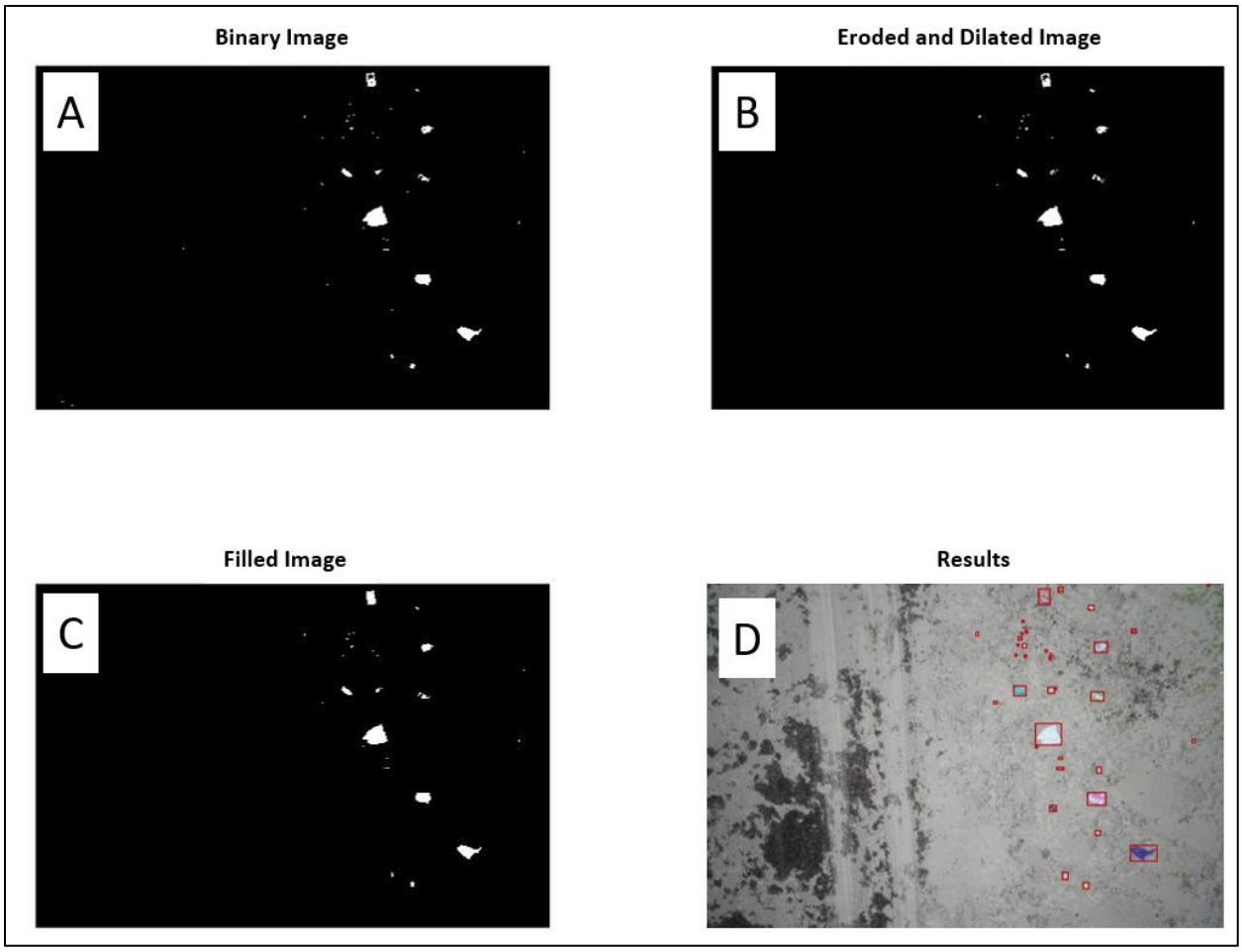


Figure 14: Morphological processing: (A) binary image includes the noise of background, (B) and (C) the image after applying the dilation, erosion, and fill operation to (A), and the result of applying the morphological operations to detect the debris using red bounding boxes (D).

2.5 High-Level Processing

The mid-level process involves images as input with outputs being attributes extracted from those images, such as segmentation and classification. With high-level processing, image

attributes are the input, and the output is understanding, such as scene and text description and individual objects recognition. This can be achieved through deep learning; a type of machine learning that train a neural network to learn unique characteristics of the features directly from data to classify and detect specific information, in this instance, debris in aerial images. One favorite type of deep neural network is convolutional neural networks (CNN). The main advantage of using the CNN over traditional methods like segmentation is the ability to construct feature filters during the training process. With enough training data, the self-learning CNN can perform accurate image classification task and image localization. Due to the small dataset of aerial images that contains marine debris (88 images), the SEAGULL detector was unable to classify and localize the debris in the aerial images. Instead, photos taken by a smartphone were used to train the SEAGULL detector. The goal of this section is to establish a baseline for debris detection so future works can build upon it.

This training typically takes hours to days, while requiring substantial computing power. To reduce the time required for training, the Matlab General Java Heap Memory preference was set to the maximum, and the Computer Vision System Toolbox preference was set to local parallel with four workers in a parallel pool. The computer used was GeForce GTX 1080Ti with a compute capability of 6.1.

2.5.1 Transfer Learning Using AlexNet

To achieve high predictive accuracy from a CNN, the network requires large datasets to train with (Krizhevsky et al., 2012). A dataset of that size was not feasible for this study. Therefore, this project used transfer learning. This involves fitting the debris and non-debris datasets to an existing pre-trained network. AlexNet has been trained with over a million images using the ImageNet database and can classify objects found in the image into ~1000 categories (Krizhevsky, Sutskever,

& Hinton, 2012). Transfer learning is a process of taking a pre-trained model and adjusting the model to the custom dataset (Dan et al., 2014). The goal of this aspect of the project was to adapt a pre-trained convolutional neural network (AlexNet) for automated marine debris classification.

AlexNet won the ImageNet Large-Scale Visual Recognition Challenge in 2012, achieving the highest classification performance at that time (Han et al., 2017). The network consists of twenty-five layers: 5 convolutional layers, three fully connected layers, one image input layer, two cross-channel normalization layers, three max-pooling layers, 7 rectified linear unit layers (ReLU), two dropout layers, one softmax layer, and one classification output layer (Figure 15).

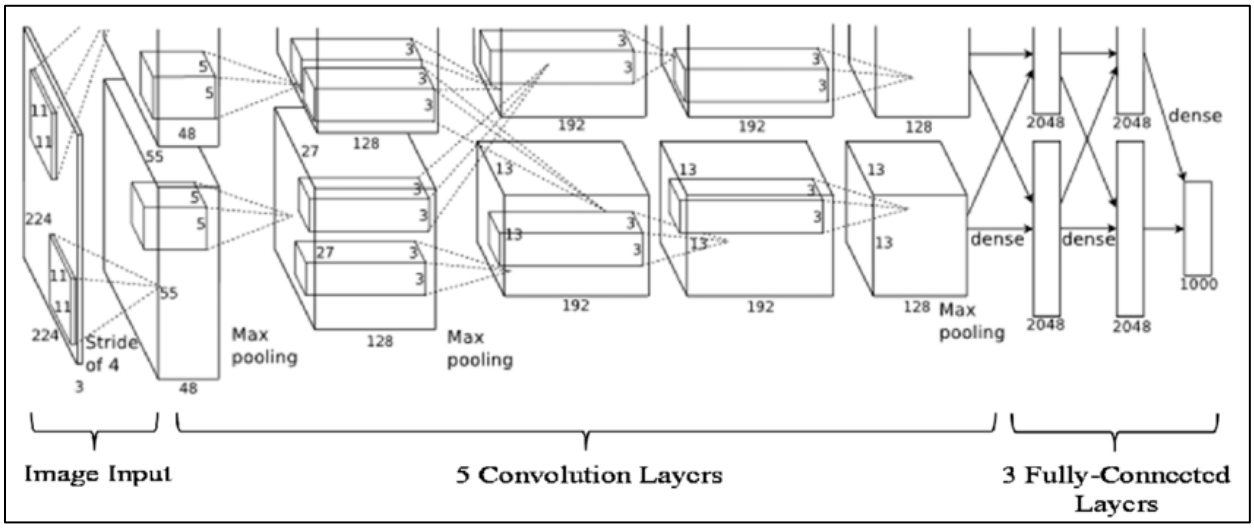


Figure 15: Architecture of AlexNet (Krizhevsky et al., 2012).

The image input layer defines the size of the input images, and AlexNet required the input image size of 227 pixels by 227 pixels with three color channels. The Debris and Non-debris datasets were resized using Matlab's Image Batch Processor application. Both datasets were randomly split into 80% for training and 20% for validation by Matlab. The training images will go through AlexNet network with the last fully connected layer being modified to train the network to classify the images into two class: Debris or Non-Debris.

The training images will go through 5 Convolution layers, each with different filter and stride number, feature maps, zero padding, output size, and neurons. The training options used for the convolutional layers was 0.001 for the initial learning rate, 64 for the mini-batch size, and 20 for the max epochs. The parameters chosen worked well with the custom datasets, and it ensures a fast training time. The initial learning rate helps to reduce the rate at which the network parameters are changing. The minibatch size is a subset of the training data that was used to evaluate the gradient of the loss function and update the weights. Lastly, the max epoch is the number of epochs used for training. Each convolutional layer has neurons or filters that learn the significant features like straight edge, colors, and curves. A ReLU layer is applied after every convolutional layer. ReLU introduces nonlinearity to the system, meaning it speeds up the training because the established threshold operation was zero. Any pixel values ≤ 0 is set to zero, and it will not propagate through the network. Due to the unbound activations of ReLU neurons, a Cross-Channel Normalization layer usually follows after the ReLU layer to normalize it. The Cross-Channel Normalization layers are channel-wise local response normalization; this layer will amplify the excited neuron while reducing the sensitivity of the neighboring neuron. The Max Pooling layers are the down-sampling layer; reducing the number of connections to the following layers and help reduce overfitting. Overfitting means that the model fits too well into the training set, making it hard to generalize to other patterns. The Fully Connected layers combine all the features learned from the previous layers to identify significant patterns. The last Fully Connected layer was modified since AlexNet was initially designed to classify images into 1000 object categories, but for this project, there were only two classes (Debris and Non-Debris). The Dropout layers set the layers input elements to zero with a given probability. This layer helps with overfitting. The last two layers were Softmax and Classification, the output layers. The Softmax

Layer normalizes the output of the fully connected layers and determines how likely the image falls into each class, and the Classification Layer detects and identifies the image. After the Debris Classification Network was trained, the performance was evaluated by comparing the predicted label from the network to the testing images.

2.5.2 Datasets

CNN's advantage is that once trained with existing datasets, it automates feature extraction, which as stated above for the mid-level processing, was initially done manually. There were two datasets for this project. The first dataset contained 1037 images of marine debris (individuals or clusters), which examples shown in Figure 16.

The second training dataset was non-debris images, which were used to reduce false positive. These images depicted naturally occurring materials such as sand, water, and coastal vegetation (some examples shown in Figure 17). Aerial images and smartphone images were used for this dataset, which contained about 1,364 images.

All of the debris dataset and some of the non-debris dataset were taken from one smartphone camera, with each image measuring ≥3000 by ≥5000 pixels. The sensor width of the camera is ~5.92mm, and the focal length of the camera is ~ 27mm. Pictures were taken between ~ 1-1.5m altitude to produce a GSD of ~0.01cm/pixel.

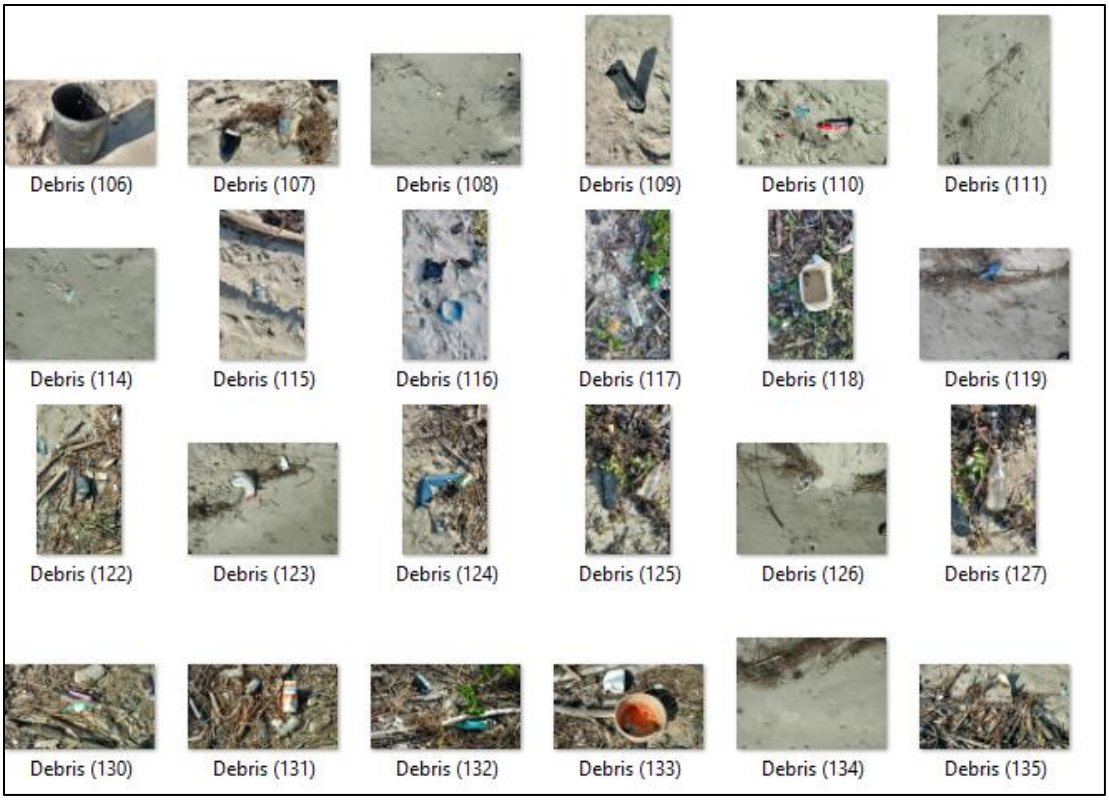


Figure 16: Samples from the Debris Dataset.

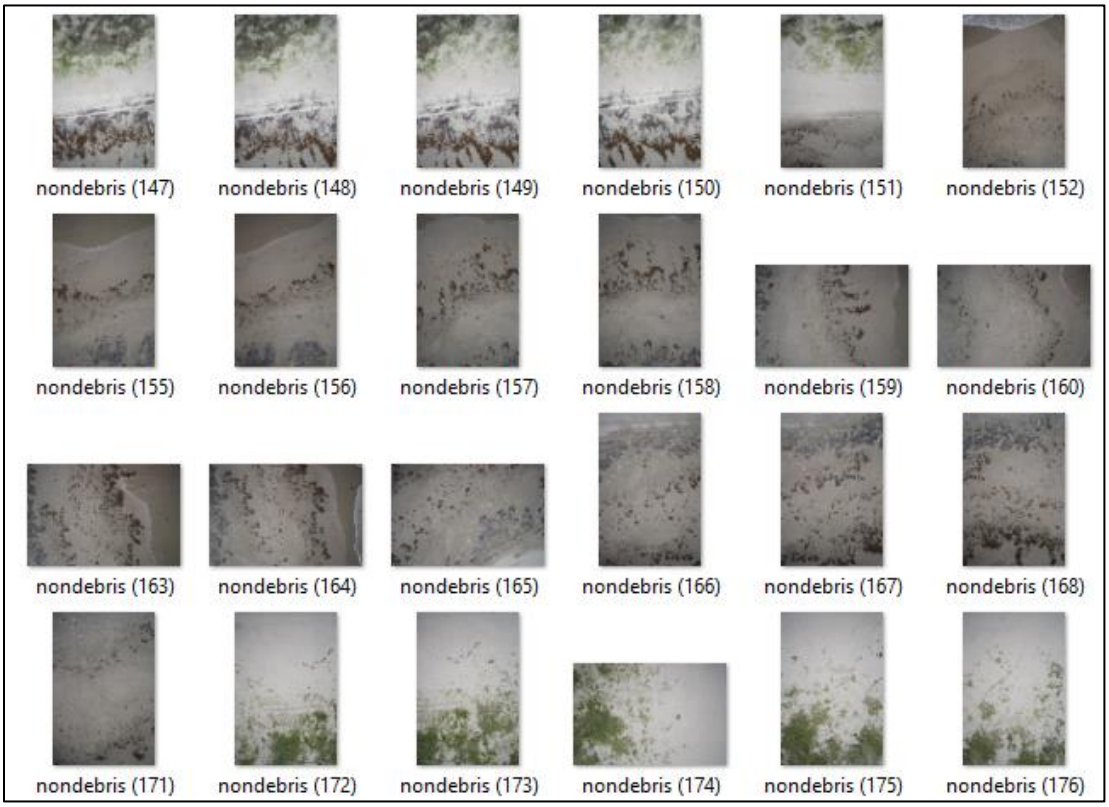


Figure 17: Samples of the Non-debris dataset.

To adapt the pre-trained neural network, AlexNet, both datasets (Debris and Non-Debris) were used for object classification training. The adapted network was further modified with R-CNN. This was done by further processing the debris image dataset using the Matlab Image Labeler application to identify debris and label it as a rectangular region of interest (ROIs; Figure 18). No cluster images (Figure 19) or those with coarse sand and white shells background (Figure 20) was used to train the R-CNN because it would increase the false positive and the false negative rate. After labeling, 839 images were exported to Matlab for training and testing the R-CNN.

Figure 18: Each image contained ROIs for debris detection.

Figure 19: An image containing a variety of debris.

Figure 20: A blue plastic object in a coarse sand environment with the white shells.

2.5.3 Region with CNN (R-CNN) Training Parameters

The goal of any object detector is to find the location of objects in an image and classify them correctly (Ibrahim et al., 2012). After image classification training was accomplished using AlexNet, R-CNN was trained to locate all relevant objects. Matlab R-CNN has three stages: extract the region proposals, train a neural network to classify objects in the training data, and train the bounding box regression models for each object class (Figure 21). After inputting the images, R-CNN creates bounding boxes or region proposals through a selective search to capture all possible object locations (Figure 21B). The selective search looks at the image through different size windows and will group adjacent pixels by texture, color, or intensity. Once the proposals are created, R-CNN warps the regions to a standard square size and passes it through the modified AlexNet (Figure 21C). The last layer will use SVM to classify the object and its location (Figure 21D). The bounding boxes will run through a linear regression model to output tighter coordinates for the box once the object is classified.

Matlab has a "trainRCNNObjectDetector" function that requires ground truth training data, a network (AlexNet), and pre-determined training options. Due to the massive raw image size, the debris dataset was resized to maintain its existing high resolution. Resizing the images also reduced Matlab memory usage and increased the performance speed of R-CNN training. The images were all resized with a width and height of 666 pixels and 1000 pixels. These values were chosen because the 2.54cm diameter debris pieces were still recognizable, and it contained enough pixels to extract relevant information (Figure 22).

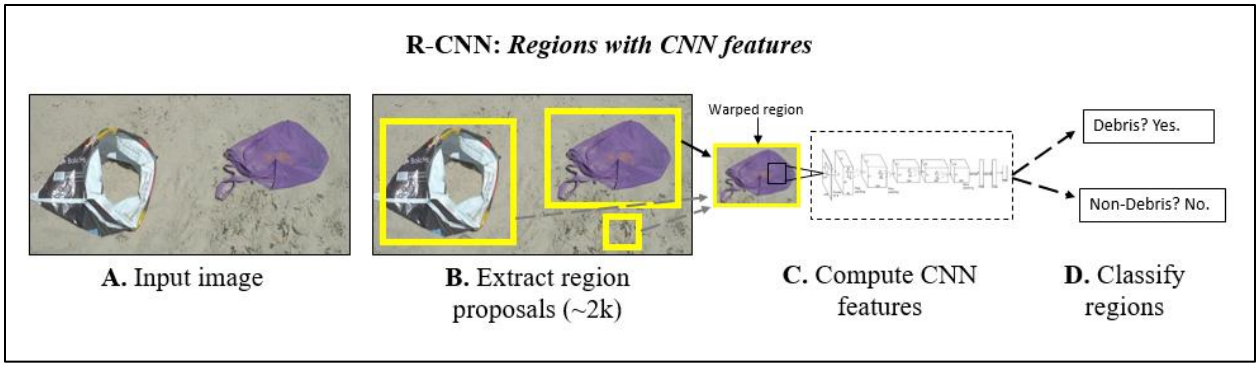


Figure 21: A modified overview of R-CNN for marine debris detection (original R-CNN figure is from Girshick et al., 2016).

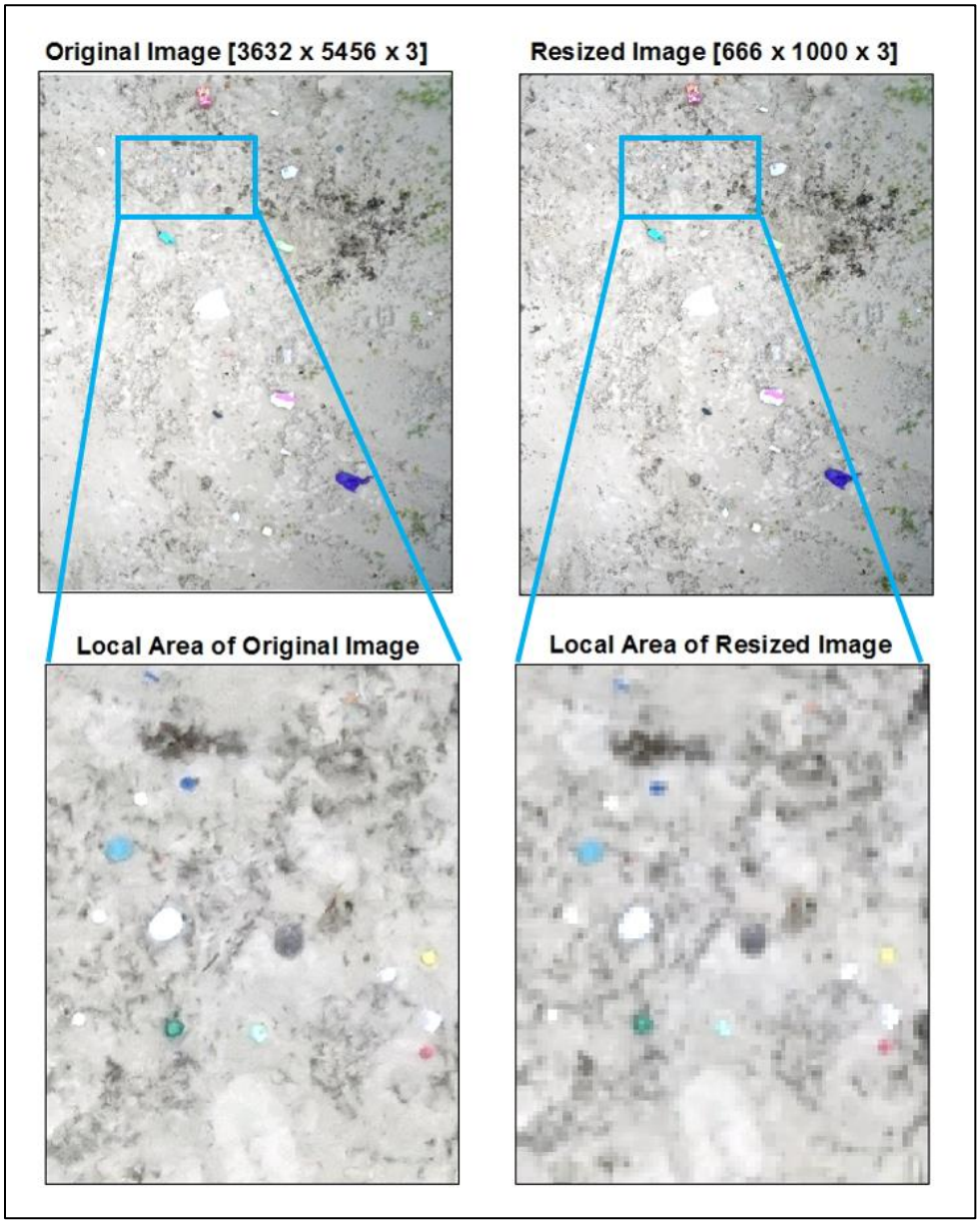


Figure 22: Resolution comparison of the original image vs. the resized image.

To start training the R-CNN for debris detection in the images, a ground truth data table was exported from the Image Labeler Application. This table contains ROI labels [x, y, width, height], which are bounding boxes that specify the object position in each image. The ground truth table was randomly split with 70% for training dataset and 30% for the testing dataset. The network used for R-CNN was the modified AlexNet that classified images into debris or non-debris class. The training options for R-CNN was set to use the stochastic gradient descent with momentum optimizer. This helped to accelerate gradient vectors in the right direction and learn the optimal filters for object detection. The training option also set the mini batch size to 96, the initial learning rate to 0.001, and the max epochs to 32. Training the R-CNN took a couple of hours.

To use the SEAGULL detector, the "detect" function in Matlab is used on a test image. The location of the objects detected and the detection scores for each bounding box are calculated. Detection scores range from 0-1 with larger values indicative of higher detection confidence. An annotation function was used to display the image with the bounding boxes and scores (Figure 23). The SEAGULL detector was tested against three different backgrounds (fine sand, coarse sand, and carpet) and a couple of aerial images at 8m and 10m.

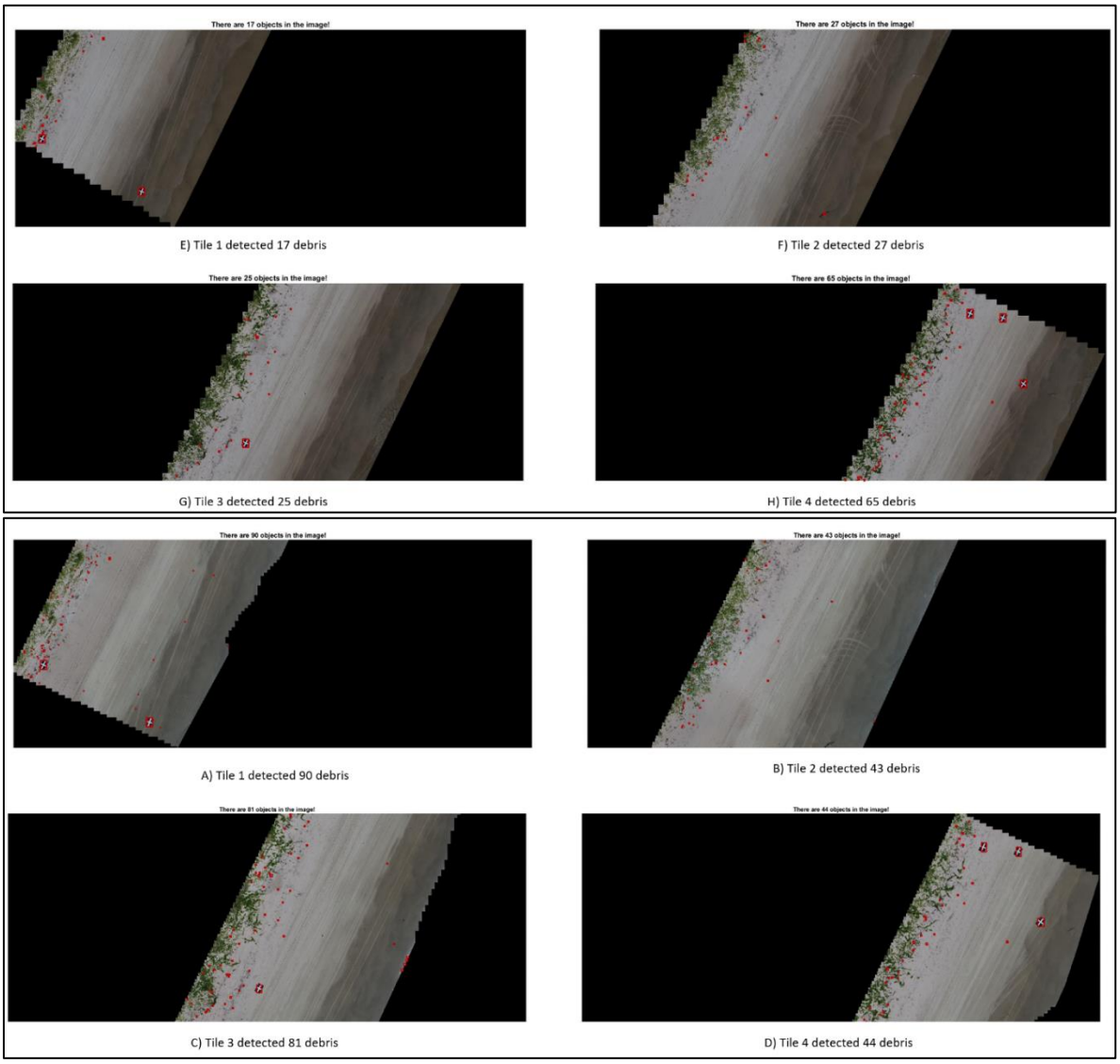
The SEAGULL detector was evaluated using the "evaluateDetectionPrecision" function (Average Precision metric). The average precision provides a single number that incorporates the ability of the detector to make the correct classification (precision) and the ability of the detector to find all relevant objects (recall). The first step is to run the detector on the test dataset (30% of the total dataset). The Average Precision metric was used to evaluate the expected results vs. the results from the detector, and a precision/recall curve is generated to visualize the relationship between precision and recall. Ideally, the precision should be one at all recall levels. If the average precision is less than 0.5, additional training data and time may be required.

Figure 23: An example of R-CNN object detection.

CHAPTER III: RESULTS AND EVALUATION

3.1 Segmentation by Color Threshold

The amount of marine debris detected through segmentation varies due to weather condition, especially the amount of sunlight reflecting off the debris, and the threshold range required to remove the majority of the background. During the August 2017 shoreline survey, a total of 341 macro-debris (>2.5cm diameter) items were manually recorded. The segmentation algorithm detected a total of 258, 134 and 57 debris items at 15, 22, and 35m. The debris detected is identified in Figure 24 by red bounding boxes.



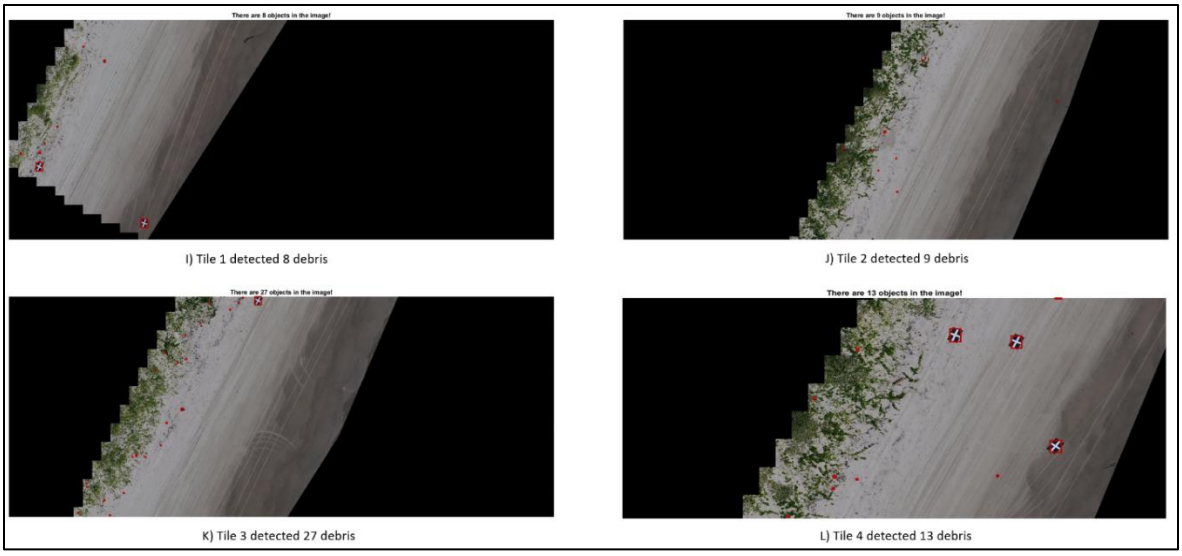


Figure 24: Result of detection for marine debris at 15m (A-D), 22m (E-H), and 35m (I-L).

However, the true positive (actual debris after removing false positives) for 15, 22, and 35m was 215, 101 and 50 debris items. Separating the foreground and background with three channels was challenging, and each flight had false positives and false negatives (Table 2). Examples of false positives included flowers or water reflection and the false negatives were attributed to debris having similar pixel values to sand and vegetation (Figure 25). These items include cardboard, transparent glass or plastic bottles, and black objects. Additionally, multicolored objects were recognized by the algorithm as multiple items (Figure 26). Overall, the detection rate at 15, 22 and 35m was 63.0, 29.6, and 14.7% over a 100m section of beach. While most of the color debris was detected at all altitudes, fragmented and therefore smaller debris (<2.5cm diameter) was not detected by the algorithm in images captured at higher altitudes.

Table 2: Quantity and assessment of debris detected at various altitudes for August 2017.

Altitude (m)	Objects Identified (out of 341)	True Positive	False Positive	False Negative	Detection Rate
15	258	215	43	126	63.0%
22	134	101	33	240	29.6%
35	57	50	7	291	14.7%

Figure 25: Two objects were missed due to the similar pixel values as sand (yellow circles).

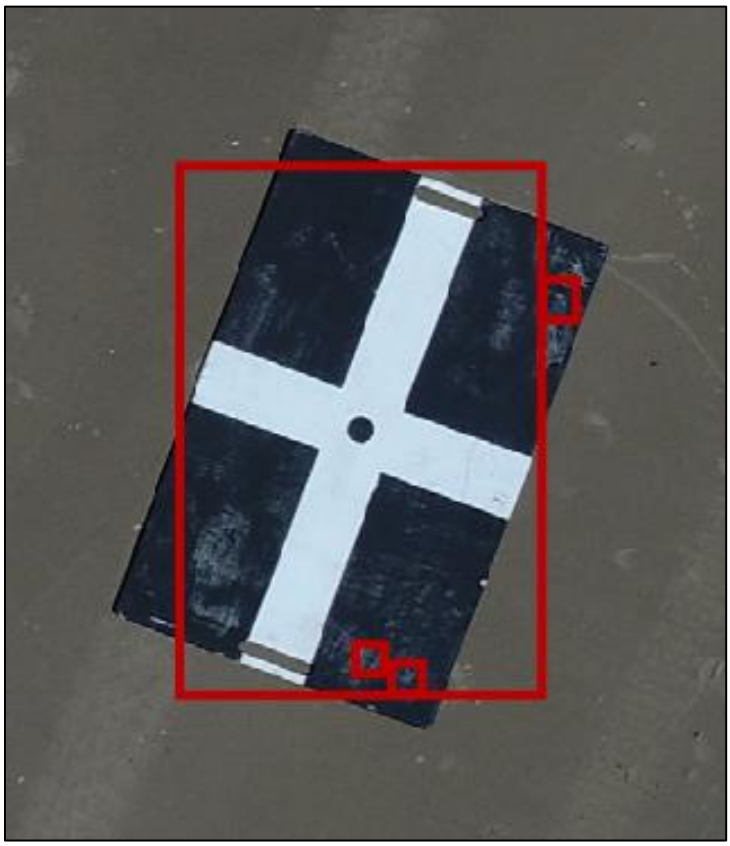


Figure 26: The algorithm recognized this target as multiple bits of debris even though it was only one object.

3.2 Object Detection by R-CNN

With object detection, it was evaluated based on determining the existence (classification) and location (localization) of an object in an image.

3.2.1 AlexNet Results and Accuracy

AlexNet was modified to sort images into two classes: Debris or Non-Debris. There were 1,037 images for debris and 1,364 images for non-debris for a total of 2,401. These images were pooled, and Matlab randomly split the two datasets where 80% (1,921 images) were used for training, while 20% (480) were used for testing. Correctly classifying the images is part of the R-CNN evaluation, so more images were put into the training process to have a higher precision. AlexNet produced 99.58% accuracy with the testing dataset. As shown in Figure 27, AlexNet was able to classify the images with high confidence.

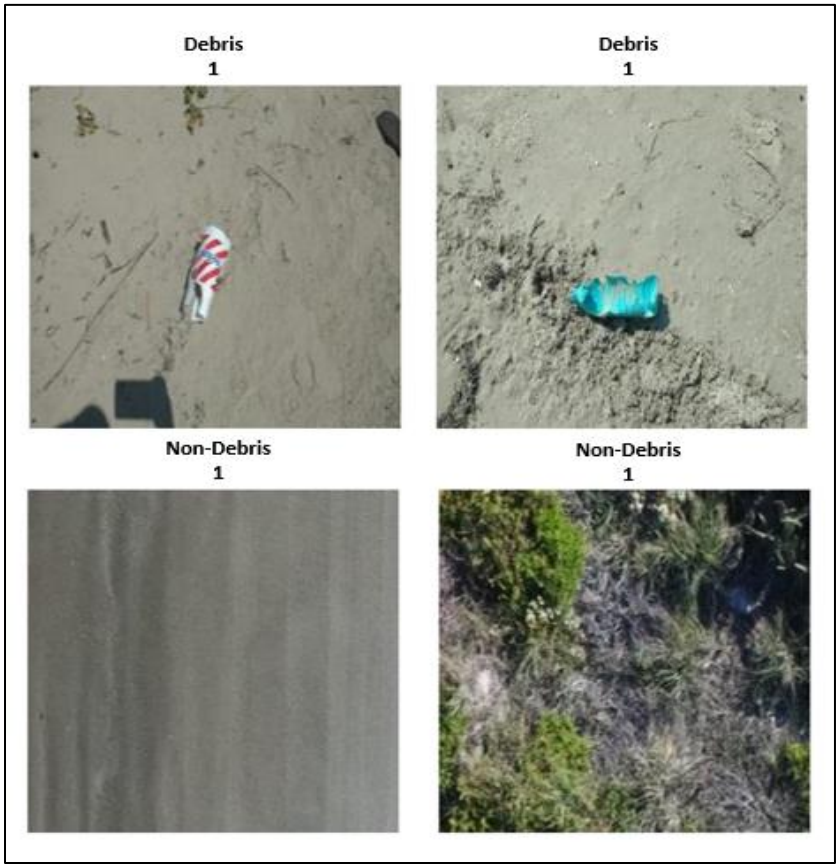


Figure 27: Each image contained a confidence score below the classification.

3.2.2 R-CNN Results and Average Precision

There are three stages in the training process of R-CNN in Matlab extract region proposals, training a neural network to classify objects in training data (AlexNet) and training bounding box regression models for each object class (Girshick et al., 2016).

The debris dataset used for R-CNN contained 839 labeled images, with 251 images (30%) for testing and 539 images (70%) for training. The size of the training dataset was smaller than the original debris dataset because cluster images and coarse sand with white shells background were removed to decrease false positive and false negative detection. The ground truth labels for the debris images were used to extract positive region proposals, and the negative regions came from the background of the image. A positive overlap range was included in the training to ensure that the region proposals that overlap the ground truth bounding box within the range of [0.6 1] were used as positive training samples. Through Selective Search, R-CNN will come up with ~2000 fixed size regions per image that possibly contain the objects. There were some images where R-CNN could not extract the region proposal even with a ground truth label (Figure 28). This may be due to background color relative to the debris or its size. It also did not process an image with a light blue plastic film and a green bottle cap. Interestingly, while the ground truth labels of debris found in a coarse sand background were rejected like Figure 28, similar debris in a fine sand background was accepted.

Figure 28: This training image was rejected because the detector cannot extract any region proposal.

The training bounding box regression models for each object class was programmed by Matlab to improve localization performance. A linear regression model was used to refine the bounding box coordinates, and it applied a Greedy Non-Maximum Suppression (Greedy NMS) to eliminate duplicate detections and to select the best scoring window; assuming that it covers an object. However, there are some downsides when using a Greedy NMS; it accepts the top score box even though it may not be the best fit, it will suppress nearby objects, and it does not suppress false positives. In Figure 29, the red spoon was partially detected, and it might be because the other proposed bounding box was not higher than 0.87177, so it was removed.

Figure 29: Plastic spoons were detected with a bounding box (yellow) and a confidence score using a greedy NMS.

Once trained, the average precision was calculated by comparing the ground truth data to the detection results. Precision looks at how many predicting items were relevant objects and recall represents how many relevant objects were predicted. The average precision for R-CNN SEAGULL detector was 21.86% (Figure 30). The low average precision for the SEAGULL detector may have been due to a high false positive and false negative rate. For example, using two images from the testing dataset for R-CNN, the image with ground truth labels shown in Figure 31A identifies only one object, but the SEAGULL detector identified five additional false positives (Figure 31B). The SEAGULL detector was also tested against a sand and vegetation background (Figure 31C) with six known items, but the SEAGULL detector only identified one, resulting in five false negatives (Figure 31D). These issues might be the result of the small debris dataset used

to train the model and having three separate stages does not allow any update to network layers when training.

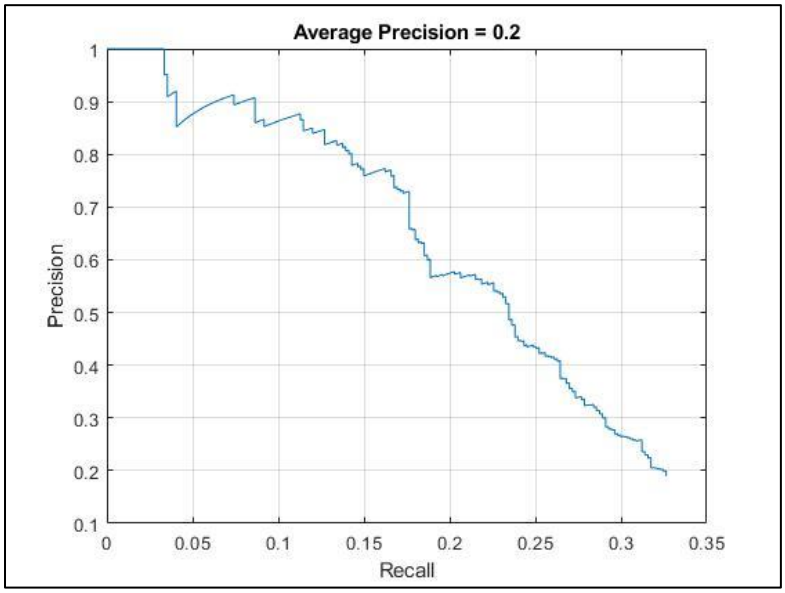


Figure 30: Precision-Recall curve for the SEAGULL detector.

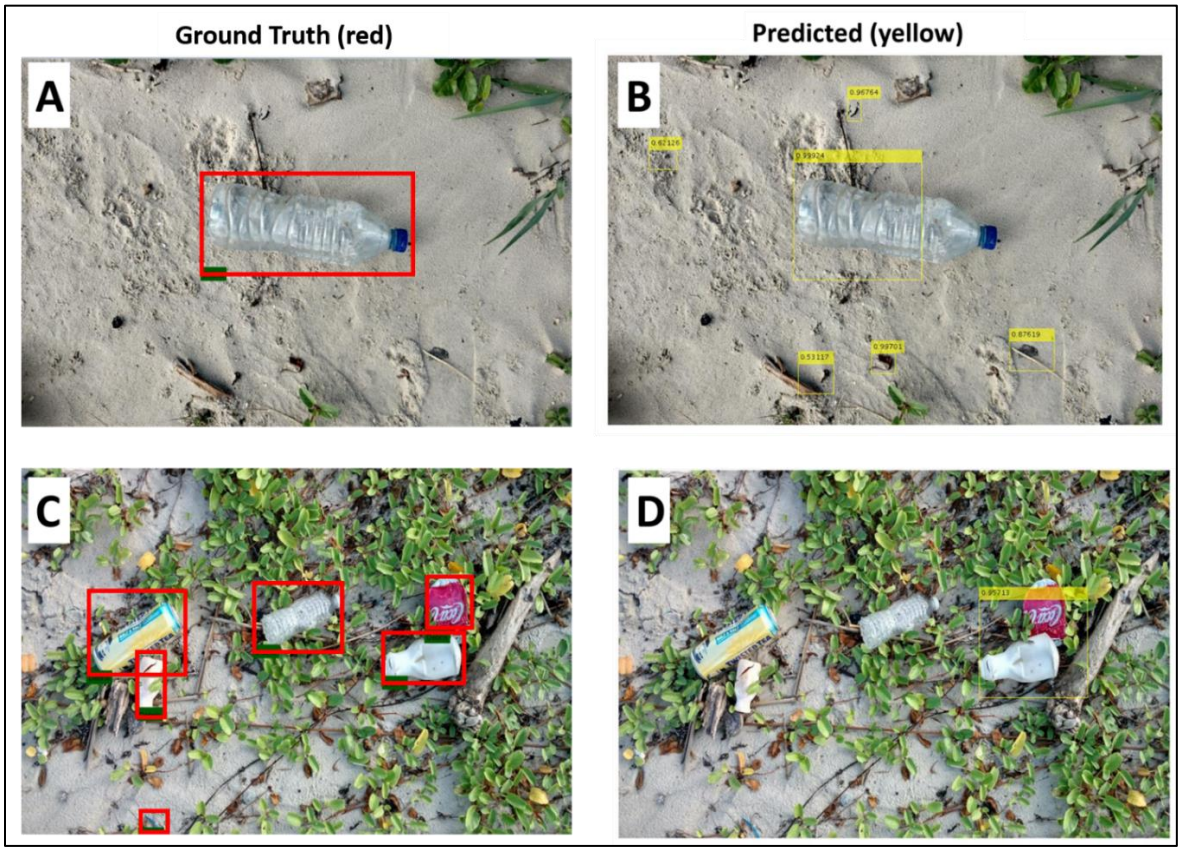


Figure 31: Ground truth bounding boxes (A & C, red) compared with the predicted bounding boxes and their confidence score (B& D, yellow).

The SEAGULL detector was tested against three different backgrounds using a smartphone camera (fine sand, coarse sand, and carpet) and three aerial images taken with the Phantom 4 Pro (0.21cm/pixel and 0.27cm/pixel). With the fine sand, the SEAGULL detector located seven out of ten debris items. Of those seven debris items, the confidence score ranged from 72%-100% for debris classification (Figure 32). It missed two clear plastic items (cap and bottle), a green plastic tube, and part of a buried red plastic spoon. However, with the coarse sand, the SEAGULL detector was only able to detect one out of nine objects (Figure 33). The SEAGULL detector was also tested using a carpeted background; it was able to locate six out of six objects with confidence score from 64%-100% (Figure 34). Even though the SEAGULL detector was able to locate and classify all the objects, it was not able to adequately locate the black fork. This may be an issue similar to Figure 29 that is related to greedy NMS.

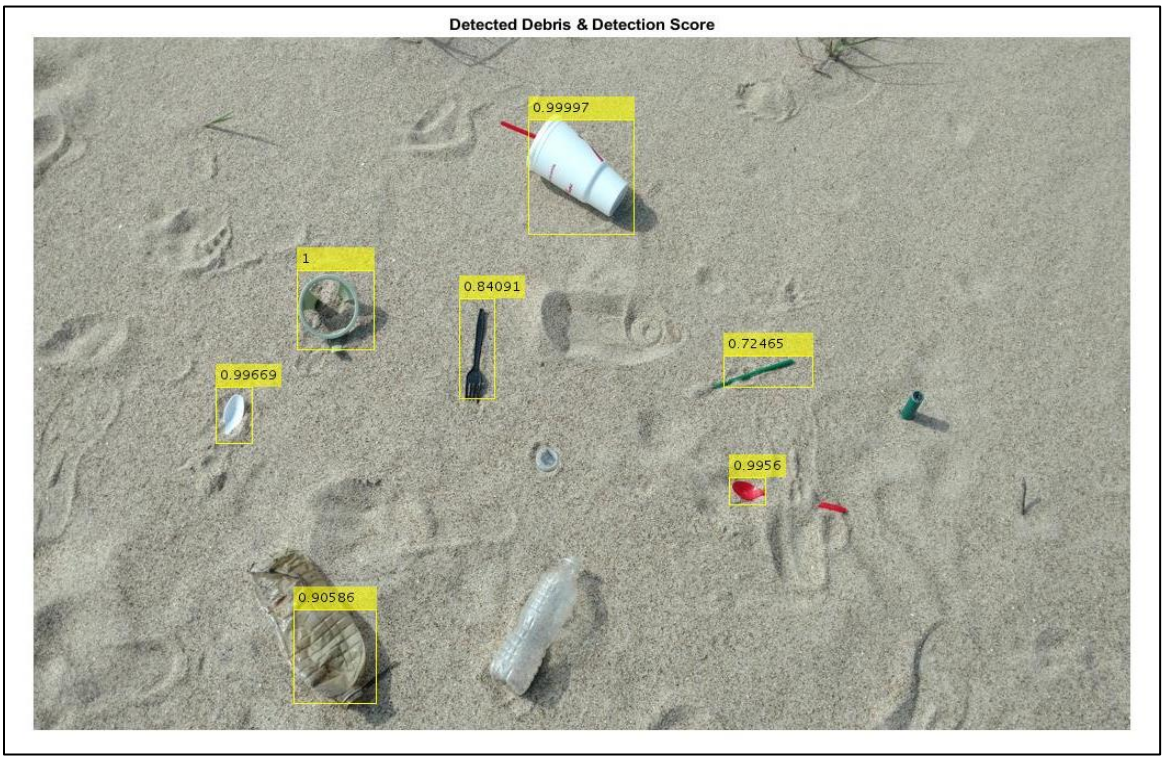


Figure 32: Detected debris in a fine sand background with a bounding box (yellow) and confidence score.

Figure 33: Detected debris in a coarse sand background with a bounding box (yellow) and confidence score.

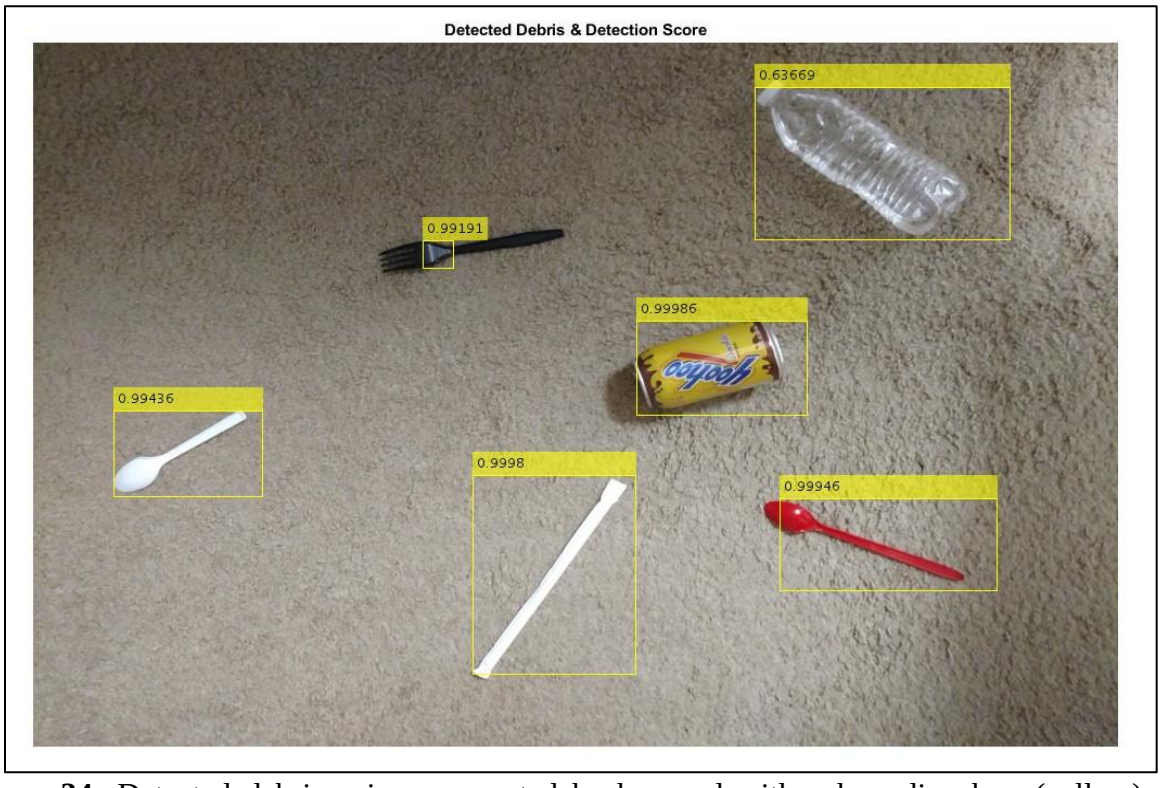


Figure 34: Detected debris using a carpeted background with a bounding box (yellow) and confidence score.

Due to the minimal aerial dataset that contains marine debris, it was not used to train the R-CNN. However, the SEAGULL detector was used for the 8m and 10m aerial images that contained staged debris. The SEAGULL detector identified five out of 12 objects for the 8m with one misclassification (Figure 35). Another 8m aerial image was tested, and the SEAGULL detector located four out of four debris but also falsely identified four vegetation items (Figure 36). In the 10m aerial image, 10 out of 36 objects were detected (Figure 37). The false negatives were likely missed because their size was less than the smallest object used during training. Overall, using deep learning based on visual detection of marine debris on the beach is plausible, considering the SEAGULL detector was capable of understanding and recognizing different types of debris.

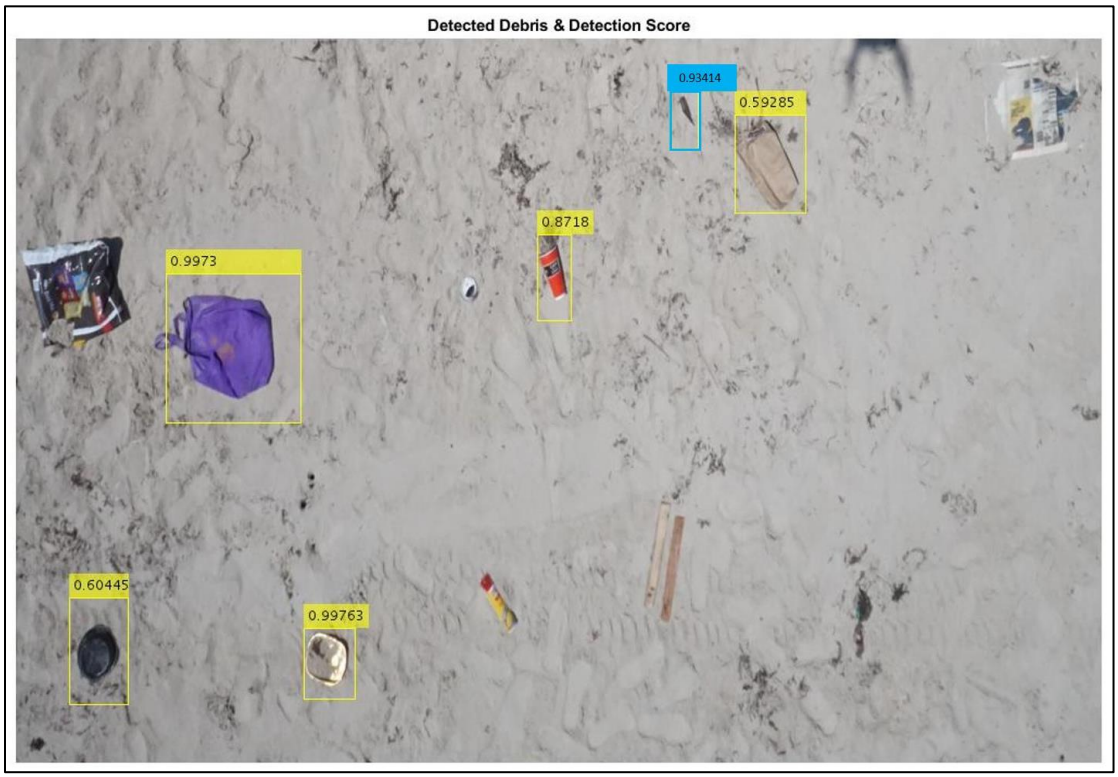


Figure 35: The SEAGULL detector located 5 out of 12 debris (yellow box) using the aerial image taken at 8m; with one misclassification (blue box).

Figure 36: The SEAGULL detector located four out of four debris (yellow box) using the aerial image taken at 8m. However, the detector also misclassified vegetation (blue box) as debris.

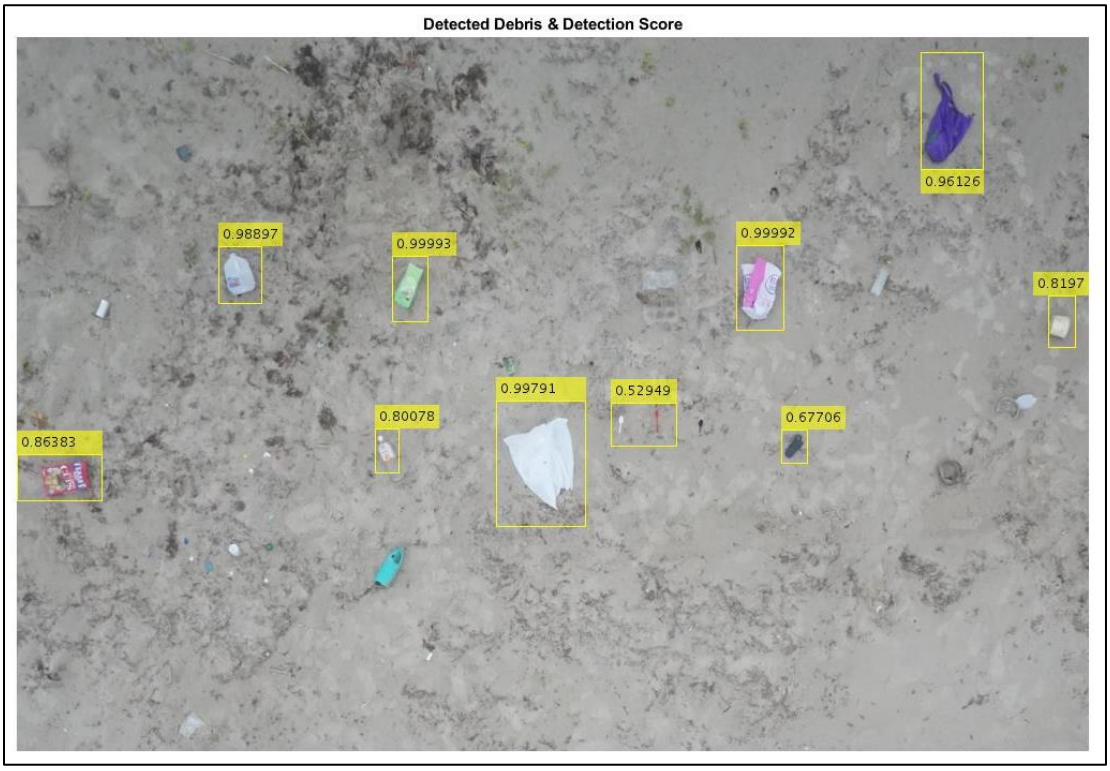


Figure 37: The SEAGULL detector located 10 out of 36 debris (yellow box) using the aerial image taken at 10m.

CHAPTER IV: DISCUSSION AND CONCLUSION

4.1 Discussion

Marine debris presents physical, biological, and chemical threats to coastal and marine ecosystems (Gall & Thompson, 2015). These threats can compromise the balance of marine ecosystems, resulting in costly control efforts, cleanups, and adverse economic impacts (Mouat et al., 2010). Beach surveys typically involve volunteer groups collecting and characterizing debris abundance, but these surveys are limited due to a significant amount of labor required. UAS beach surveys require only 1-2 people, cover large swatch of beach in a relatively short amount of time and can be performed as needed with minimal planning. This research used high-resolution UAS aerial imagery to established two image processing methods to detect and quantify marine debris in a coastal environment.

The methods developed and applied in this study to advance the ability of researchers to systematically quantify marine debris over an expansive coastline in several ways. UAS can collect data in a shorter time frame than a traditional beach survey, obtain high-resolution imagery, and allowed for repeated analysis when necessary. Essentially it can overcome the spatial and temporal limitations of traditional beach surveys. However, UAS operation is limited by weather and FAA restrictions. Jang et al. (2014) developed a similar technique for detecting beach litter using UAS: converted RGB to HSI color space, created a binary image, performed morphological operations on the image, and extracted the debris. However, Jang et al. (2014) only flew the UAS at ~10m and only used debris larger than 50cm in diameter because their Beach Litter Detection software was tested with 50cm by 50cm markers. This research was able to assess the impact of flight design and ground sample distance on detection and quantification of marine debris as varies altitudes. From this project, the maximum altitude for flying the UAS should be 35m. While 35m altitude

may not be suitable for detecting the small debris (bottle caps and fragmented plastics), it might be suitable for coastal environments with large items such as buoys, ghost nets, and vessels. Through segmentation by color thresholding, the most accurate flight elevation tested was 15m.The lower the altitude, the higher the resolution, allowing fragmented debris to be detected better. However, with top-down photography, only the superficial layer of debris was detected which can result in an underestimate of the total debris on the beach. The spatial coverage of marine debris was not calculated because the debris pixels were altered during the Morphological Operation stage when the background pixels were removed. Also, objects with similar pixel values as sand and vegetation were either ignored or misclassified as a non-debris item. This problem has also been encountered by Nakashima et al., 2011; Kataoka et al., 2012; and Jang et al., 2014, which used a different remote sensing technique or color space and they still could not detect the clear plastics or objects with similar pixel values as sand and vegetation. Then again, they all used an RGB sensor so that future work could incorporate a multispectral or hyperspectral imaging camera on the UAS. These cameras use visible, near-infrared (NIR), and shortwave infrared wavebands when capturing images. This additional information, including the brightness and spectral reflectance, could be useful for object detection. Goddijn-Murphy et al. (2018) developed an optical reflectance model for measuring light reflectance off of floating macro-plastics. While this model is only applied to a single smooth flat layer, it does take into consideration the color, transparency, reflectivity, and shape of the plastic debris. Since beached debris is usually biofouled by bacteria, algae or barnacles, this can reduce the reflectance of visible light. However, since algae contain chlorophyll, NIR would be useful to look at the high reflectance from the algae biofilm to detect the debris (Goddijn-Murphy et al., 2018). Other components that could affect the spectral reflectance of macro-plastics is the color, condition (dry or wet), and polymer type (Garaba

& Dierssen, 2018). Another example is to use NIR and red band to create a vegetation index, which can be used to exclude vegetation from the orthomosaic. Attempts like this at reducing background noise associated with non-target pixels would increase the amount of debris detected and reduce the false positives. Multispectral sensors can also be flown at night to determine differences in radiant flux between the plastic and its surroundings. The debris density, thermal conductivity, and specific heat might be different from sand, and the different spectral channels can help in debris detection. However, for researchers with limited access to advanced sensors and software, manual tagging of debris on the orthomosaic is acceptable to determine location and abundance as well as some physical characteristics (Moy *et al*. 2017). Even though it can be time-consuming and the results may vary from observer to observer, visual observation can help reduce the false positive detection and missed objects.

4.2 Conclusion

With recent advancement in deep learning based computer vision model, object detection applications have been well researched in many fields. CNN have revolutionized object detection and recognition, but with regards to marine debris, this deep-learning approach has only been applied to submerged materials captured with an autonomous underwater vehicle (Valdenegro-Toro, 2016; Fulton et al., 2018). This research, however, looked at beached debris and created two custom datasets to train an R-CNN. Even though the dataset to train the R-CNN was small compared to using the ImageNet database, the accuracy for debris classification through AlexNet was ~99%. However, the average precision for debris detection through R-CNN was 22%. Even though the average precision is low, the initial results are promising. The overall low performance of the SEAGULL detector might be due to the small amount of training data, the complex background used to test R-CNN, and 25 layers used for debris classification. In the future, we hope

to expand upon this work by increasing the marine debris dataset, but also test it with similar algorithms such as Fast R-CNN and Faster R-CNN. These updated versions of R-CNN should increase the overall accuracy, have faster processing time, and the ability to re-train the different stages to decrease false positive and false negative rates. By increasing the marine debris database, we hope to classify and detect the different types of debris (plastic, metal, paper, wood, and other types of debris.) on the beach. Also, by collaborating with other researchers, we can look at the benefits and limitations of using different object detector networks such as You Look Only Once (YOLO), YOLO version 2, and Single Shot MultiBox Detector.

These findings can help establish a structured protocol for mapping and monitoring debris along with any coastline, especially remote or inaccessible area where beach surveys are not an option. The capability for rapid deployment, survey collection, and data processing makes UAS an optimal survey tool to conduct repeated surveys over areas of limited spatial coverage. Incorporating UAS surveys into the beach litter monitoring program can "help identify, determine sources of, assess, reduce, and prevent marine debris and its adverse impacts on the marine environment (Marine Debris Research, Prevention, and Reduction Act, 33 USC §§ 1951)." The methods developed and applied in this research lay a foundation that can be used to vastly improve our ability to study the fate and transport of marine debris in coastal systems.

REFERENCES

1. Belhouari, Y., Farnum, B., Jenkins, C., Kieser, J., Lopez de Roman, A., McCauley, D., Rochman, C., Schreiber, R., Schwartz, E., Taylor, H., Trott, S., 2017. International Coastal Cleanup 2017 Report.
2. Browne, M.A., Galloway, T., Thompson, R., 2007. Microplastic-An Emerging Contaminant of Potential Concern? Integr. Environ. Assess. Manag. 3, 559–566. https://doi.org/10.1897/IEAM_2008-022.1
3. Browne, M.A., Chapman, M.G., Thompson, R.C., Amaral Zettler, L.A., Jambeck, J., Mallos, N.J., 2015. Spatial and Temporal Patterns of Stranded Intertidal Marine Debris: Is There a Picture of Global Change? Environ. Sci. Technol. 49, 7082–7094. https://doi.org/10.1021/es5060572
4. Chen, Z., Zhang, T., Ouyang, C., 2018. End-to-End Airplane Detection Using Transfer Learning in Remote Sensing Images. Remote Sens. 10, 139. https://doi.org/10.3390/rs10010139
5. Chudasama, D., Patel, T., Joshi, S., Prajapati, G.I., 2015. Image Segmentation using Morphological Operations. Int. J. Comput. Appl. 117, 16–19.
6. Critchell, K., Grech, A., Schlaefer, J., Andutta, F.P., Lambrechts, J., Wolanski, E., Hamann, M., 2015. Modelling the fate of marine debris along a complex shoreline: Lessons from the Great Barrier Reef. Estuar. Coast. Shelf Sci. 167, 414–426. https://doi.org/10.1016/j.ecss.2015.10.018
7. Corcoran, P.L., Biesinger, M.C., Grifi, M., 2009. Plastics and beaches: A degrading relationship. Mar. Pollut. Bull. 58, 80–84. https://doi.org/10.1016/j.marpolbul.2008.08.022
8. Dan, Z., Sang, N., He, Y., Sun, S., 2014. An improved LBP transfer learning for remote

sensing object recognition. Optik (Stuttg). 125, 482–485. https://doi.org/10.1016/j.ijleo.2013.07.021

9. Derraik, J.G.B., 2002. The pollution of the marine environment by plastic debris: A review. Mar. Pollut. Bull. 44, 842–852. https://doi.org/10.1016/S0025-326X(02)00220-5
10. Duhec, A. V., Jeanne, R.F., Maximenko, N., Hafner, J., 2015. Composition and potential origin of marine debris stranded in the Western Indian Ocean on remote Alphonse Island, Seychelles. Mar. Pollut. Bull. 96, 76–86. https://doi.org/10.1016/j.marpolbul.2015.05.042
11. Engler, R.E., 2012. The complex interaction between marine debris and toxic chemicals in the ocean. Environ. Sci. Technol. 46, 12302–12315. https://doi.org/10.1021/es3027105
12. Fulton, M., Hong, J., Sattar, J., 2018. Robotic Detection of Marine Litter Using Deep Visual Detection Models. IEEE/RSJ Int. Conf. Intell. Robot. Syst.
13. Galgani, F., Hanke, G., Maes, T., 2015. Global Distribution, Composition, and Abundance of Marine Litter, Marine Anthropogenic Litter. https://doi.org/10.1007/978-3-319-16510-3
14. Gall, S.C., Thompson, R.C., 2015. The impact of debris on marine life. Mar. Pollut. Bull. 92, 170–179. https://doi.org/10.1016/j.marpolbul.2014.12.041
15. Garaba, S.P., Dierssen, H.M., 2018. An airborne remote sensing case study of synthetic hydrocarbon detection using short wave infrared absorption features identified from marine-harvested macro- and microplastics. Remote Sens. Environ. 205, 224–235. https://doi.org/10.1016/j.rse.2017.11.023
16. Ge, Z., Shi, H., Mei, X., Dai, Z., Li, D., 2016. Semi-automatic recognition of marine debris on beaches. Sci. Rep. 6, 25759. https://doi.org/10.1038/srep25759
17. Girshick, R., Donahue, J., Darrell, T., Malik, J., 2016. Region-Based Convolutional

Networks for Accurate Object Detection and Segmentation. 142 IEEE Trans. PATTERN Anal. Mach. Intell. 38, 142–158.

18. Goddijn-Murphy, L., Peters, S., van Sebille, E., James, N.A., Gibb, S., 2018. Concept for a hyperspectral remote sensing algorithm for floating marine macro plastics. Mar. Pollut. Bull. 126, 255–262. https://doi.org/10.1016/j.marpolbul.2017.11.011
19. Gonzalez, R.C., Woods, R.E., 2007. Digital Image Processing (3rd Edition), Prentice-Hall, Inc. Upper Saddle River, NJ, USA ©2006. https://doi.org/10.1017/CBO9781107415324.004
20. Gregory, M.R., 2009. Environmental implications of plastic debris in marine settings--entanglement, ingestion, smothering, hangers-on, hitch-hiking and alien invasions. Philos. Trans. R. Soc. B Biol. Sci. 364, 2013–2025. https://doi.org/10.1098/rstb.2008.0265
21. Han, X., Zhong, Y., Cao, L., Zhang, L., 2017. Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification. Remote Sens. 9. https://doi.org/10.3390/rs9080848
22. Ibrahim, M.S., Badr, A.A., Abdallah, M.R., Eissa, I.F., 2012. Bounding box object localization based on image superpixelization. Procedia Comput. Sci. 13, 108–119. https://doi.org/10.1016/j.procs.2012.09.119
23. Itkin, M., Kim, M., Park, Y., 2016. Development of cloud-based UAV monitoring and management system. Sensors 16, 1–19. https://doi.org/10.3390/s16111913
24. Jambeck, J.R., Roland, G., Wilcox, C., Siegler, T.R., Perryman, M., Andrady, A., Narayan, R., Law, K.L., 2015. Plastic waste inputs from land into the ocean, Science. https://doi.org/10.1017/CBO9781107415386.010
25. Jang, S.W., Kim, D.H., Chung, Y.H., Seong, K.T., Yoon, H., 2014. Development and

application of a technique for detecting beach litter using a Micro-Unmanned Aerial Vehicle. Korean J. Remote Sens. 30, 351–366.

26. Jang, S.W., Lee, S.K., Kim, D.H., Chung, Y.H., Yoon, H.J., 2015. Application of Remote Environmental Monitoring Technique to Efficient Management of Beach Litter. Int. J. u-e- Serv. Sci. Technol. 8, 357–368.
27. Kako, S., Isobe, A., Magome, S., 2010. Sequential monitoring of beach litter using webcams. Mar. Pollut. Bull. 60, 775–779. https://doi.org/10.1016/j.marpolbul.2010.03.009
28. Kataoka, T., Hinata, H., Kako, S., 2012. A new technique for detecting colored macro plastic debris on beaches using webcam images and CIELUV. Mar. Pollut. Bull. 64, 1829–1836. https://doi.org/10.1016/j.marpolbul.2012.06.006
29. Kosuth, M., Mason, S.A., Wattenberg, E. V, 2018. Anthropogenic contamination of tap water , beer , and sea salt 1–18. https://doi.org/10.7910/DVN/IFCKDL.Funding
30. Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks. Adv. Neural Inf. Process. Syst. 1–9. https://doi.org/http://dx.doi.org/10.1016/j.protcy.2014.09.007
31. Lee, J., Hong, S., Song, Y.K., Hong, S.H., Jang, Y.C., Jang, M., Heo, N.W., Han, G.M., Lee, M.J., Kang, D., Shim, W.J., 2013. Relationships among the abundances of plastic debris in different size classes on beaches in South Korea. Mar. Pollut. Bull. 77, 349–354. https://doi.org/10.1016/j.marpolbul.2013.08.013
32. Lippiatt, S., Opfer, S., Arthur, C., 2013. Marine Debris Monitoring and Assessment. NOAA Tech. Memo. NOS-OR&R-46.
33. Marine Debris Research, Prevention, and Reduction Act, 2006. 33 U.S.C. §§ 1951 et seq
34. Mathews, A.J., Jensen, J.L.R., 2013. Visualizing and quantifying vineyard canopy LAI

using an unmanned aerial vehicle (UAV) collected high density structure from motion point cloud. Remote Sens. 5, 2164–2183. https://doi.org/10.3390/rs5052164

35. Mouat, J., Lozano, R.L., Bateson, H., 2010. Economic Impacts of Marine Litter. Kommunernes Int. Mijloorganisation 117.
36. Moy, K., Neilson, B., Chung, A., Meadows, A., Castrence, M., Ambagis, S., Davidson, K., 2017. Mapping coastal marine debris using aerial imagery and spatial analysis ☆. Mar. Pollut. Bull. https://doi.org/10.1016/j.marpolbul.2017.11.045
37. Nakashima, E., Isobe, A., Magome, S., Kako, S., Deki, N., 2011. Using aerial photography and in situ measurements to estimate the quantity of macro-litter on beaches. Mar. Pollut. Bull. 62, 762–769. https://doi.org/10.1016/j.marpolbul.2011.01.006
38. National Park Service, 2005. Padre Island: Administrative history. Chapter Eight: Natural resource issues. Published June 2005 <https://www.nps.gov/parkhistory/online_books/pais/adhi8.htm>.
39. Opfer, S., Arthur, C., Lippiatt, S., 2012. NOAA Marine Debris Shoreline Survey Field Guide. NOAA Mar. Debris Program. 19.
40. Pan, S.J., Yang, Q., 2010. A survey on transfer learning. IEEE Trans. Knowl. Data Eng. 22, 1345–1359. https://doi.org/10.1109/TKDE.2009.191
41. Phantom 4 Pro, n.d. <https://www.dji.com/phantom-4-pro>
42. Ramesh, K.N., Chandrika, N., Omkar, S.N., Meenavathi, M.B., Rekha, V., 2016. Detection of Rows in Agricultural Crop Images Acquired by Remote Sensing from a UAV. Int. J. Image, Graph. Signal Process. 8, 25–31. https://doi.org/10.5815/ijigsp.2016.11.04
43. Recreational Users, 2015. < http://knowbeforeyoufly.org/for-recreational-users/>.
44. Ribic, C.A., Dixon, T.R., Vining, I., 1992. Marine debris survey manual. Mar. Pollut. Bull.

26, 348. https://doi.org/10.1016/0025-326X(93)90583-6

45. Ribic, C.A., Sheavly, S.B., Rugg, D.J., 2011. Trends in Marine Debris in the U.S. Caribbean and the Gulf of Mexico 1996-2003. J. Integr. Coast. Zo. Manag. 11, 7–19. https://doi.org/10.5894/rgci181
46. Seymour, A.C., Dale, J., Hammill, M., Halpin, P.N., Johnston, D.W., 2017. Automated detection and enumeration of marine wildlife using unmanned aircraft systems (UAS) and thermal imagery. Sci. Rep. 7, 1–10. https://doi.org/10.1038/srep45127
47. Sun, T., Wang, Y., Yang, J., Hu, X., Member, S., 2017. Convolution Neural Networks with Two Pathways for Image Style Recognition. IEEE Trans. Image Process. 26, 4102–4113. https://doi.org/10.1109/TIP.2017.2710631
48. Thiel, M., Hinojosa, I.A., Miranda, L., Pantoja, J.F., Rivadeneira, M.M., Vásquez, N., 2013. Anthropogenic marine debris in the coastal environment: A multi-year comparison between coastal waters and local shores. Mar. Pollut. Bull. 71, 307–316. https://doi.org/10.1016/j.marpolbul.2013.01.005
49. Turner, I.L., Harley, M.D., Drummond, C.D., 2016. UAVs for coastal surveying. Coast. Eng. 114, 19–24. https://doi.org/10.1016/j.coastaleng.2016.03.011
50. Valdenegro-Toro, M., 2016. Submerged marine debris detection with autonomous underwater vehicles. Int. Conf. Robot. Autom. Humanit. Appl. RAHA 2016 - Conf. Proc. https://doi.org/10.1109/RAHA.2016.7931907
51. Veenstra, T.S., Churnside, J.H., 2012. Airborne sensors for detecting large marine debris at sea. Mar. Pollut. Bull. 65, 63–68. https://doi.org/10.1016/j.marpolbul.2010.11.018
52. Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., Reynolds, J.M., 2012. “Structure-from-Motion” photogrammetry: A low-cost, effective tool for geoscience

applications. Geomorphology 179, 300–314. https://doi.org/10.1016/j.geomorph.2012.08.021

53. Xu, F.-H., Chang, Y.-L., Oey, L.-Y., Hamilton, P., 2013. Loop Current Growth and Eddy Shedding Using Models and Observations: Analyses of the July 2011 Eddy-Shedding Event*. J. Phys. Oceanogr. 43, 1015–1027. https://doi.org/10.1175/JPO-D-12-0138.1
54. Yang, C.Z., Yaniger, S.I., Jordan, V.C., Klein, D.J., Bittner, G.D., 2011. Most plastic products release estrogenic chemicals: A potential health problem that can be solved. Environ. Health Perspect. 119, 989–996. https://doi.org/10.1289/ehp.1003220