NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

EFFECTS OF TARGET CLASSIFICATION ON AI-BASED UNEXPLODED ORDNANCE DETECTION PERFORMANCE

by

Haocheng Joel Li

September 2021

Thesis Advisor: Oleg A. Yakimenko
Second Reader: Fotis A. Papoulias

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2021	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE EFFECTS OF TARGET CLASSIFICATION ON AI-BASED UNEXPLODED ORDNANCE DETECTION PERFORMANCE	**5. FUNDING NUMBERS**
6. AUTHOR(S) Haocheng Joel Li	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

This thesis aims to reduce the safety risks for warfighters in an area of operations where unexploded ordnance (UXO) may be present, and lessen the number of training opportunities due to malfunctioning munitions in a controlled environment. The thesis leverages the advancement in unmanned technologies and artificial intelligence (AI) development to complete dull, dirty, and dangerous tasks more effectively. Specifically, the thesis attempts to improve a trained AI detector's performance using different data-labeling methods as applied to the electro-optical images. The thesis describes the efforts conducted to train a UXO detector for a proposed deep learning convolutional neural network followed by validating its performance. To further enhance UXO detection capabilities, the research explores how the optimal target classification method developed and verified for a single-spectrum sensor can also be applied for a multispectral sensor. As such, the thesis outlines a development of a prototype of a real-time UXO detection system composed of a commercial-off-the-shelf (COTS) multi-spectral sensor and a small COTS unmanned aerial system.

14. SUBJECT TERMS small unmanned aerial system, UXO detection, deep learning, neural network , target classification, target labelling	**15. NUMBER OF PAGES** 71
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

EFFECTS OF TARGET CLASSIFICATION ON AI-BASED UNEXPLODED ORDNANCE DETECTION PERFORMANCE

Haocheng Joel Li
Military Expert 5, Singapore Army
BME, Nanyang Technological University, 2013

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
September 2021

Approved by: Oleg A. Yakimenko
Advisor

Fotis A. Papoulias
Second Reader

Oleg A. Yakimenko
Chair, Department of Systems Engineering

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This thesis aims to reduce the safety risks for warfighters in an area of operations where unexploded ordnance (UXO) may be present, and lessen the number of training opportunities due to malfunctioning munitions in a controlled environment. The thesis leverages the advancement in unmanned technologies and artificial intelligence (AI) development to complete dull, dirty, and dangerous tasks more effectively. Specifically, the thesis attempts to improve a trained AI detector's performance using different data-labeling methods as applied to the electro-optical images. The thesis describes the efforts conducted to train a UXO detector for a proposed deep learning convolutional neural network followed by validating its performance. To further enhance UXO detection capabilities, the research explores how the optimal target classification method developed and verified for a single-spectrum sensor can also be applied for a multispectral sensor. As such, the thesis outlines a development of a prototype of a real-time UXO detection system composed of a commercial-off-the-shelf (COTS) multi-spectral sensor and a small COTS unmanned aerial system.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AI	artificial intelligence
AP	average precision
CNN	convolutional neural network
DL	deep learning
DLCNN	deep learning convolutional neural network
DLS	Downwelling Light Sensor
DOD	Department of Defense
EO	electro-optical
IoU	Intersection over Union
ML	machine learning
MS	multi-spectral
NDAA	National Defense Authorization Act
PR	precision-recall
RGB	Red-Green-Blue
ROI	region of interest
RPN	region proposal network
SSD	Single Shot MultiBox Detector
sUAS	small unmanned aerial system
UAS	unmanned aerial system
UXO	unexploded ordnance
YOLO	You-Only-Look-Once

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to take this opportunity to thank my thesis advisor, Professor Oleg Yakimenko, for his unwavering support and overwhelming understanding during the thesis journey. Professor Yakimenko has been extremely generous in imparting valuable knowledge and sharing countless advice on the subject matter. I could not have completed my thesis without him. Our interactions and discussions sparked my interest in computer vision and model-based systems engineering. I also wish to thank the faculty of the Systems Engineering Department for their assistance and professionalism rendered to me during my research at the Naval Postgraduate School. I would like to give my heartfelt thanks to CPT Cho Seungwan for his generous sharing of tacit knowledge on the subject matter and friendship. Most importantly, I want to express my gratitude to my beloved wife, Ting Rui, for her care and understanding throughout this period.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Unexploded ordnance (UXO) poses a serious threat to military troops and assets operating in a hostile area of operations. It also contributes to lost training opportunities due to the occurrence of munition malfunctions in a controlled environment. Presently, UXO detection operations require intensive manpower and time resources (Etter and Delaney 2013). UXO clearance operations are also high-risk activities that expose personnel to catastrophic consequences. Capitalizing on technological advancement and maturity in the areas of unmanned aerial systems (UAS) and object detection using electro-optical (EO) sensors, coupled with deep learning (DL) algorithms, this thesis explores different data processing factors when employing an AI-based detection algorithm from data collected by small unmanned aerial systems (sUAS).

A. BACKGROUND

With the understanding of the limitations and risks involved in the current concept of UXO detection and capitalizing on the opportunities that emerging technologies offer, the value proposition of this thesis is to (1) reduce the safety risks posed to civilians and military troops in areas that contain UXO by increasing detection probability, and (2) lessen the resources needed to conduct UXO clearing operations by decreasing the time required to complete the detection operations.

1. UXO Landscape

Philip Grone (2003) defined a UXO as military-grade munitions that 1) have been "primed, fused, armed, or prepared for" usage, 2) were projected, dropped, launched, delivered, or located to create a "hazard to operations, installations, personnel," and 3) remained "unexploded due to design, malfunction or any causes" (4). Figure 1 shows some examples of possible UXO. As mentioned previously, the current method and process of UXO detection has safety and throughput limitations. Usually, UXO require surface level (non-invasive) detection that uses visual or optical scanning. At times, UXO can be sub-surfaced, requiring magnetometer, active electromagnetic induction, and ground penetrating radar detectors to locate such UXO (Bertrand et al. 2004). The equipment used

in such instances is usually manpower intensive as it is operated by handheld means and requires personnel to perform physical scanning work to locate the presence of any UXO. This method inherently creates operational risks to the operator.

Figure 1. Examples of UXO. Source: Australia DOD (n.d.).

Furthermore, UXO decontamination programs also entail high financial costs. The "21 mine action projects listed in the U.N. Mine Action Strategy's portfolio cost an average of US $27 million each" (Carter Center 2020, 11). Therefore, increasing the probability of detection through tasks automation will increase the efficiency of performing area decontamination and reduce the cost of such operations.

In a land-scarce country like Singapore, the availability of a live-firing area meant for training is crucial for soldiers to achieve mastery of weaponry competency and combat units to maintain warfighting capabilities. Therefore, it is paramount to reduce the downtime necessary to remedy UXO incidents in the designated impact zone during training. This, in turn, will allow training sequences to resume quickly and achieve training objectives in their designated timeframe.

2. UXO Threats in Area of Military Operation

A 2020 Carter Center study estimated a minimum of 94,792 individual explosive munitions were used during armed conflicts between government forces and local militias

within a radius of 105 square kilometers of the Syrian capital of Damascus between July 2013 and May 2019 (Carter Center 2020). In fact, about 5 percent of all munitions being used in any modern conflict can potentially become UXO (Overton. 2020). These estimates highlight the potential extent of UXO presence within a small, populated area and the resulting danger UXO can pose after a conflict. The UXO scenario not only creates safety issues during periods of armed conflict but the contamination from the remaining explosive content of the UXO weapons will continue to harm people for an extended period after an armed conflict (Carter Center 2020).

These risks from UXO threaten military personnel and assets operating in the area of operations. As UXO are often undetected after armed kinetic conflict, personnel and assets might not have prior intelligence about these risks during follow-on operations in the area. This UXO threat can hinder the tempo of military operations and consequently reduce the military efforts' effectiveness and cause injury and even death when personnel unknowingly come across UXO.

Civilians and their way of life are also affected by UXO, even after a conflict is over (Carter Center 2020). Many areas previously used by the military may remain under-utilized or not used at all because of UXO threats. There are many instances where civilians' livelihoods depend heavily on agriculture, which is particularly vulnerable to dangers posed by remaining UXO and can deny civilians access to large sections of land. Figure 2 shows members of the Afghan National Army searching a road for improvised explosive devices before marking the area safe for travel and operations.

Figure 2. Improvised Explosive Device Clearance. Source: Savell (2019).

3. Opportunities from Advancements in Technology

In recent times, development of unmanned technologies has provided promising results for unmanned systems to undertake dangerous, dull, dirty, and dangerous tasks instead of human operators (Bekmezci et al. 2016). Artificial Intelligence (AI) has also created opportunities in military applications by applying DL algorithms based on artificial neural networks.

a. Small Unmanned Aircraft System

The U.S. Department of Defense (DOD) has defined Group 1 UAS as "typically less than 20 pounds in weight and normally operate below 1200 feet above ground level at speeds less than 250 knots" (U.S. DOD 2010, 12). This classification falls within the Federal Aviation Administration's definition of any unmanned aircraft or aerial vehicle weighing less than 55 pounds.

The U.S. Army has indicated that it currently employs UAS at different echelons of its force structure, and in particular, it will continue to equip Battalion–level and below units with sUAS (Group 1) as an organic asset for tactical operations. These sUAS could be deployed for full-spectrum military operations and to provide "just-in-time" situational awareness (U.S. DOD 2001). The U.S. DOD has also begun integrating UAS into its force structure more widely, and these integration efforts highlight the importance of unmanned

systems in future warfare. Specifically, the U.S. DOD aims to achieve autonomy and human-machine collaboration to revolutionize warfighting concepts (U.S. DOD 2018).

It is evident that sUAS will be more widely used for military operations in the future just by noting the rising percentage of sUAS inventory of the U.S. Army. A 2013 Department of Transportation Report presented that there are about 6,200 sUAS in the U.S. Army's aircraft fleet, which is equivalent to about 55 percent of the fleet. This percentage is expected to increase to more than 75 percent, or about 10,000 sUAS, by 2035 (U.S. Department of Transportation 2013).

It was envisaged that the U.S. Army would continue to fully integrate sUAS into its concept of operations during the period of 2016 to 2035 (U.S. DOD 2010). Figure 3 depicts the U.S. Army's prediction of work-share between unmanned and manned systems, in which UAS will cover a large majority of surveillance and communication tasks and about half of the offensive mission profile from 2016 to 2035.

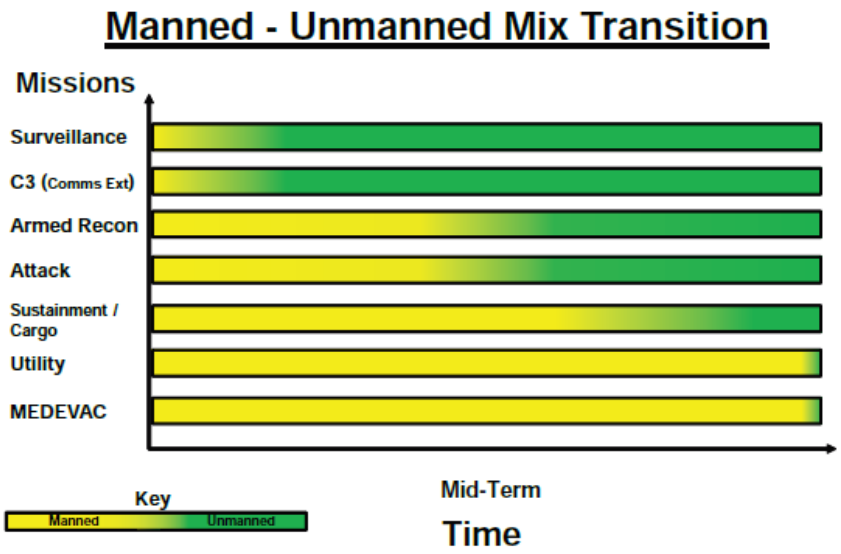


Figure 3. Envisaged Manned-Unmanned Roles Transition. Source: U.S. DOD (2010).

These unmanned systems are projected to have a full suite of detection capabilities such as optical sensors ready to be deployed for imagery capturing capabilities, including

visible light, multi-spectral (MS), or thermal images. With the improved functionality and benefits, the employment of sUAS could be applied to improve the current methodology of UXO detection. As a result, the safety of personnel could be improved.

b. Object Recognition Technologies

AI has become prevalent in our everyday lives and is defined by Vibhuthi Viswanathan (2018) as intelligence demonstrated by computers or machines. Machine learning (ML) is a type of AI technology that involves machines that can self-learn and improve by reviewing large data sets without intervention from humans. Viswanathan (2018) explained that ML can be further broken down into the category of DL which mimics the way the human brain gains knowledge. He further added that DL employs artificial neural networks for processing data and aids in decision making. The learning from DL is derived from observing large data sets that pass through artificial networks and are used for recognizing patterns and classifying objects. The availability of the internet has facilitated more effective deep learning as users can now access huge datasets easily (Khan and Salim 2020).

In a blog post, Jason Brownlee (2019) has described object recognition as a combination of "related computer vision tasks that involve identifying objects in digital images." He further explained that these tasks include image classification that predicts the type and location of objects within an image or video with a bounding box. Computer vision uses pattern recognition and image mapping to make sense of data and arrive at solutions (Esposito and Donato 2001). This technology considers an image to be an array of pixels and automates tasks by employing ML techniques (Bond et al. 2019). Advancements in object detection technology such as the use of a trained deep learning convolutional neural network (DLCNN) during object recognition tasks have allowed such tasks to be carried out more precisely (Liu and Lang 2019).

There are a variety of sensor technologies available commercially to collect data for and perform object detection. For example, typical electro-optical (EO) sensors combine Red-Green-Blue (RGB) spectrum images using a single lens, and MS sensors combine multiple lenses on a single machine to capture images of different sets of

wavelength bands to generate adequate spectral contrast within a single global shutter activity. The MS sensors can capture wavelengths invisible to the human eye, and these MS images can used to detect disturbed soil features caused by changes in the physical environment and in chemical properties. Such soil features can be used to locate UXO more effectively. MS images can also distinguish the difference in the reflectance of objects and background to display more separation, which is useful during UXO classification and detection.

With all these considerations, this author understood the practicality of performing UXO recognition using DLCNN from MS images captured from an sUAS.

B. PROBLEM FORMULATION AND THESIS ORGANIZATION

The threats posed by the UXO and the opportunities created by recent technological advancements in AI technologies inspired the idea of AI-based UXO detection using sUAS equipped with optical sensors. It is envisioned to make use of sUAS in UXO detection process. Furthermore, target classification, which is an integral part of UXO detection capabilities, creates the ground truth to facilitate the training and validation of the detectors. Therefore, it is paramount to understand the effects of target classification to introduce more effective UXO detection capabilities.

The objective of this thesis is to improve the present UXO detection model by employing alternative methods of EO image data processing. In particular, the thesis aims to answer the research question of how different target classifications of the EO image data will affect the performance of the UXO detector. The thesis uses EO images collected using a digital camera in a prior phase of the research as the data for training, validation, and testing of the detector. It trains the You-Only-Look-Once (YOLO) v2 neural network to perform UXO detection using different target classification methods of the EO images. Methods of classification are varied by deliberately placing tightly fitted or loosely fitted bounding boxes around the target present in the EO images data set. These respective labeled images are then used as training data to train the corresponding detectors. After completing the training, the detectors are tested by performing UXO detection by running the detector on all the images in the test dataset. An evaluation of the different models is

performed for comparison, and the variance in detection performance is attributed to the different target classifications.

In this phase of the research, images from EO sensors were used for efficiency and because of time constraints as the improved method of target classification from the EO images could be applied on the MS image dataset as well. This is possible because the same image will be captured in different spectrums when working with MS data, and while the pre-processing period would require more time resources, using an effective method to label the image datasets will allow users to reduce the time required to perform that task.

The effectiveness of UXO detection would be dependent on 1) Precision criteria, and 2) Recall criteria. Precision criteria refers to how accurate the prediction of the trained network performing the detection is; that is, the percentage of correct predictions. Recall criteria refers to the imprecision estimates of the trained network performing detection based on the omission of UXO detection. A measure of average precision (AP) would also be computed from the area under the Precision – Recall curve and used as the evaluation measure for the detection (Everingham et al. 2010). The performance measure of each of the different methods for target classification is used for comparison to identify the best data processing method for superior UXO detection capabilities.

The thesis is structured as follows: Chapter II presents the envisaged concept of operation that uses an sUAS coupled with a UXO detection system and includes the literature review. Chapter III describes the selection process of an appropriate DLCNN. The chapter also describes the process of developing the detection algorithm. Chapter IV describes the data sources and different data processing methods. Chapter V presents the training procedure of the processed EO images. The chapter also demonstrates the UXO detection and presents the evaluation of the detection algorithm. Chapter VI focuses on the integration process of the sUAS with a multispectral EO sensor prototype system. This process relied on assets available in the Naval Postgraduate School to support the ability to collect and analyze data not visible to the human eye and to accomplish these tasks at a fraction of the time and at reduced risk. Finally, Chapter VII summarizes the thesis, presents the conclusion, and proposes recommendations for future research.

II. CONCEPT OF OPERATIONS AND LITERATURE REVIEW

This chapter describes the employment of an sUAS-based UXO detection system in an operational scenario and offers a review of the literature relevant for this research and the developed system.

A. CONCEPT OF OPERATIONS

The proposed system consists of military operators performing sUAS flight during mission sets in a hostile environment. Hostile environments include not only areas of current military combat operations but also areas of previous armed conflict where UXO may remain and pose a threat. Using an sUAS equipped with optical sensors, teams could identify an ingress or egress route for troops or perform area sanitization by detecting surface UXO. The type of sensor in such a scenario can be either a standard EO or an MS sensor, which tends to provide more information for data classification on the images captured during flight.

Ground commanders would first define the search area and operate the sUAS equipped with an imagery sensor. The sUAS would perform a flight path to comb the specified area and capture video footage or series of overlapped images of the mission area. Upon returning to the take-off location, the sUAS team would perform UXO detection from the retrieved footage or series of images. The trained detector would indicate suspected UXO within bounding boxes on the video and images automatically. The location of suspected UXO would also be marked on the flight path overview for ground commanders to make informed decisions about the situation. An illustration of the system concept is provided in Figure 4.

Adapted from https://commons.wikimedia.org/wiki/File:Biskeri-_Camping_I_IMG_7238.jpg.

Figure 4. UXO Detection System Concept.

B. LITERATURE REVIEW

This section reviews materials related to this research. These topics include 1) detection using sUAS, 2) the feasibility of using DLCNN with images from sUAS, and 3) improved detection performance using MS images.

1. Object and UXO Detection Tasks Using sUAS Equipped with Different Sensors

Other than the traditional handheld UXO detection method, several methods have shown promising results for detecting objects. A thesis research in 2018 used Class 1 UAS (sUAS) for automated foreign object debris detection over a targeted area of operation. Wee Leong Lee (2018) developed algorithms to facilitate object detection and produced a graphical-user interface for demonstration of the concept. The study also showed the feasibility of achieving improved efficiency in operations and demonstrated promising results of detecting objects as small as 3 cm x 3 cm. He proposed an optimized system configuration of flight height of 4 m with an optical sensor filter window size of 7 pixels x 7 pixels. A study by Bartosz Ptak and Mateusz Piechocki in 2020 also showed affirmative

results in using aerial photos for the following tasks: 1) object detection, 2) segment tags, and 3) classifying objects. The study outlined the usefulness of ML and DL solutions for each of the subtasks. Their proposed approach performed well in stationary tests, and the authors were confident that this approach would produce similar results in the operational environment (Ptak and Piechocki 2020).

Other studies have explored alternative sUAS surface-detection methods. One such study developed a transient electromagnetic (TEM) system integrated with drones operating at low altitude to detect underground and near-surface UXO. The successful system demonstration results showed the advantages of UXO detection at low flight altitude. These advantages included low cost, increased flight safety, and increased efficiency (Qi et al. 2020).

Another research utilized drones equipped with a thermal imaging camera for rapid mine detection. Timothy deSmet et al. (2018) successfully used thermal sensing commercial-off-the-shelf UAS platforms to accurately detect minefield presence. The study also demonstrated the sUAS's capabilities of detecting UXO in various conditions that differed in temperature, moisture content, and earth-covered depth. The study further presented the benefits of using vision-based detection methods by having the ability to detect more materiel types than only metallic UXO.

These previous studies showed that the implementation of low flying sUAS equipped with sensors meant for UXO detection was feasible due to its inherent advantages like effectiveness, safety, and cost benefits.

2. Application of DLCNN with sUAS Images

Object detection technology has matured greatly in recent times by applying DL as a means of implementing CNN techniques to analyze visual imagery (Liu and Lang 2019). A study by Ross Girshick et al. (2014) showed that DLCNN was useful for delivering high performance in object detection by using extracting region proposals from an image as an input to the network. These proposals are then reshaped to a fixed size and passed on through the network. The authors further explained that these features within the proposals are subsequently used to "classify each region with a category-specific linear Support

Vector Machine" (2). The study also used the bounding box regression technique to ensure the object is properly captured by the bounding box (Girshick et al. 2014).

DLCNN has also been explored for performing inspections of concrete structures and buildings using images from sUAS. A study by Sattar Dorafshan et al. in 2018 demonstrated affirmative object detection capabilities by a network trained using images obtained from point-and-shoot camera instead of sUAS images. The study also presented that it was feasible to train DLCNN use training images datasets collected with handheld cameras and deploying sUAS autonomous structural inspections to achieve similar results to human inspectors (Dorafshan et al. 2018). This method is similarly applied in this thesis to perform network training and validation test using image datasets from handheld digital cameras.

A study in 2021 presented the feasibility of using ML to perform targeted weed control in an area of operations. The study focused on the ML techniques using data from the sUAS to produce an autonomous detection method for undesirable vegetation. The results obtained provided evidence that object detection could be achieved by using a trained convolutional neural network (CNN) with positive outcomes (White 2021).

Through capitalization of the technological advancements in data processing, object detection capabilities could be employed in military operations such as UXO detection.

3. Object Detection Using Multispectral Images

An sUAS can be equipped with many types of cameras that capture single or MS images, and these images when used as inputs for object detection have shown positive results in industrial solutions and research studies. For example, a 2009 study said that MS images obtained through sensors placed on sUAS created the capacity to meet the necessary requirements of spatial, spectral, and temporal resolutions and quick turnaround times. The use of sUAS further lowered the cost of this approach as compared to manned solutions (Berni et al. 2009). Studies have shown that, for agricultural applications, images obtained from an sUAS can achieve equal or better detection estimation as compared to manned airborne sensors. For example, spectral red-edge has proven to be highly effective

in detecting changes in vegetation because there is a linear relationship between red-edge reflectance spectra and the chlorophyll content of vegetation (Curran, Dungan, and Gholz 1990). This finding could be useful to aid UXO detection.

Furthermore, according to Delores Etter and Bill Delaney, a U.S. DOD report on UXO by the Defense Science Board Task Force, suggested that the use of MS data can contribute significant information for detection of surface or near-surface UXO (Etter and Delaney 2003). A study from Tokyo University also explored object detection using multiple spectral image data of 1) RGB, 2) Near-infrared (NIR), 3) Mid-infrared (MIR), and 4) Far-infrared (FIR). The spectral images (FIR, MIR, and NIR) have different features from RGB images and would aid in object detection. The results of the study showed the model that used multispectral images performed 13 percent above the mean average precision (AP) as compared to the model using RGB images for object detection (Takumi et al. 2017).

A recent study by Md Osman Gani et al. (2021) highlighted that MS datasets such as NIR images have shown more promising detection results during low visibility conditions. In their research UXO detection performed better when MS imagery data was used for YOLOv3 CNN training and data augmentation was done to prevent overfitting (Gani et al. 2021). As indicated by the promising results presented in these studies, multispectral data along with improved data processing, could further enhance detection capabilities in future.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DEVELOPMENT OF THE DETECTION ALGORITHM

This chapter describes the selection process of appropriate AI neural network and detector's parameters for UXO detection. This chapter also develops the UXO detection algorithm based on previous studies and introduces different data labeling methods in attempt to improve the detection capabilities of the proposed solution. In order to present a comparison of the detection capabilities, this study uses the same evaluation metric employed to evaluate detection capabilities in the previous research phase, as performed and documented by Cho (2021).

A. SELECTION OF DEVELOPMENT ENVIRONMENT

These days there are several environments (trained CNNs) that can be used for developing algorithms or programs. Some of the popular ones are Oracle, Python, C++, and Java. Each environment has its strengths and weaknesses, and the selection of the design environment should be based on the eventual operational need of the program.

This thesis research used MATLAB to perform UXO detection due to this author's familiarity with MATLAB and the availability of the programming software. MATLAB also allows the researcher to develop applications with guided user interfaces and provides associated programming toolboxes such as ML and DL toolboxes that were useful in developing the algorithm.

B. SELECTION OF DLCNN FOR UXO DETECTION

The selection of a suitable DLCNN for UXO detection was dependent on the available resources and complexity of tasks. Ozan Ozturk et. al. (2020) largely classified computer vision tasks into four groups, 1) image segmentation, 2) image classification, 3) object detection, and 4) object recognition. Image segmentation converts the image to the pixel level and segments the image into different partitions. This task could be used for object detection based on pixels (Sinha 2020). Image classification is the task of classifying the object's presence within the image as an input. Object detection is the process of finding instances of object images, whereas object recognition refers to the process of performing

object position estimation within the image input based on object classes. This task would be useful for scenarios where there are multiple classes of objects within the same image (Sharma 2019).

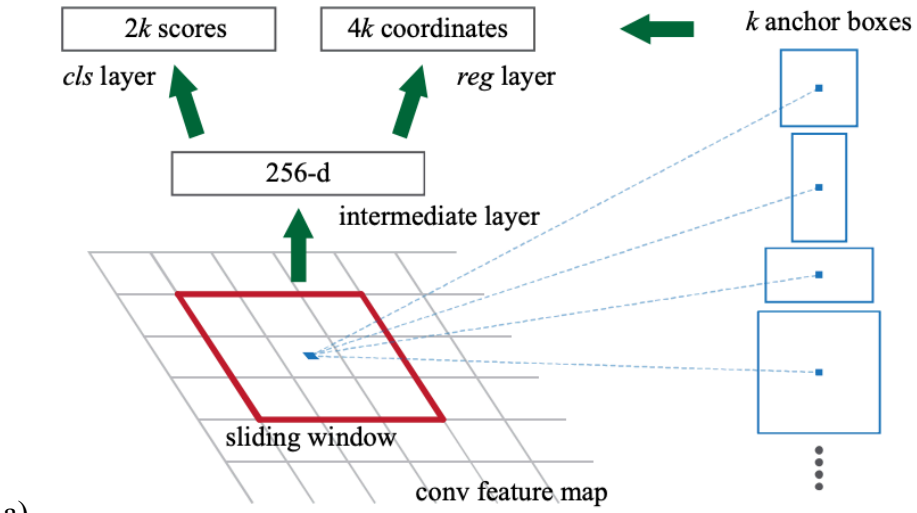
The requirements for an UXO detection system should include locality of the threat; therefore, the object detection task would employ bounding boxes to mark the predicted object in the image (Zhao et al. 2019) to warn operators of the potential danger.

The DLCNN model uses the multiple neural layers abstraction method to acquire data representations through inputs. It has proved to be more capable with an exponentially increased expressive capability (Zhao et al. 2019). Within MATLAB three popular CNN models are available, 1) YOLOv2, 2) Faster R-CNN, and 3) Single Shot MultiBox Detector (SSD).

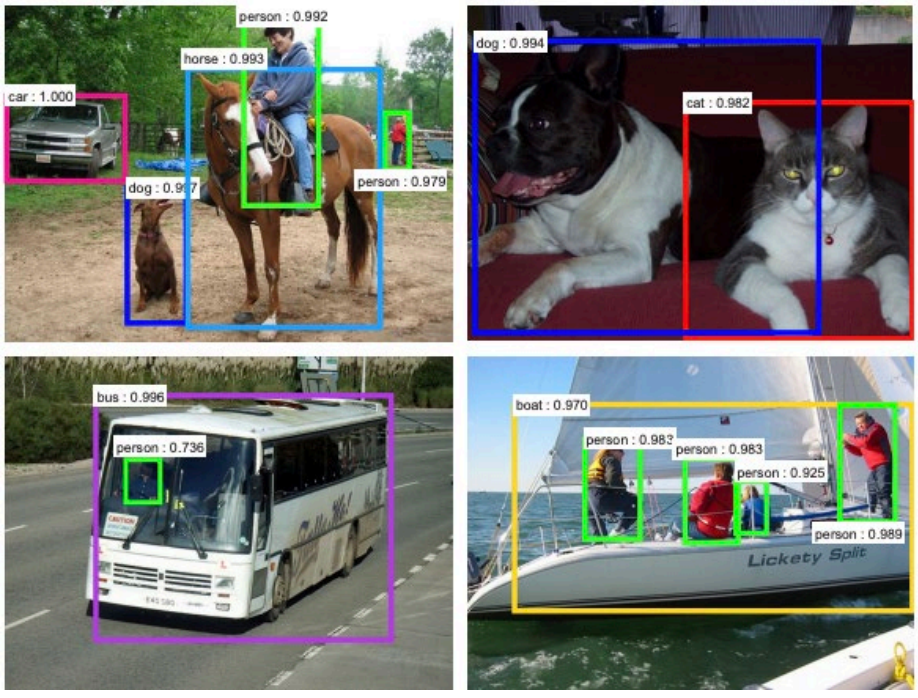
The YOLOv2 is a single-stage real-time object detection model that performs efficiently as the detection process is within a single network. It improves upon YOLOv1 in several ways, including in its use of 1) batch normalization to improve convergence as it uses higher learning rates and is less restrictive on initialization, and this feature can eliminate the need for Dropout occurrences (Ioffe and Szegedy 2015); (2) it offers a high-resolution classifier; and (3) YOLOv2 allows for anchor boxes to predict bounding boxes to frame the locality of objects presented in the image (Redmon and Farhadi 2016). The study by Joseph Redmon and Ali Farhadi (2016) also presented the YOLOv2's superior accuracy and speed in comparison to the other computational models and the preceding version—YOLO on PASCAL VOC 2007 test set—and across a variety of detection datasets.

Another model, named Faster R-CNN, developed by Shaoqing Ren et al. (2016) merged the region proposal network (RPN) and an extension of a previous network model—Fast R-CNN. The RPN helps create convolutional features in full-image format within the detection network for object bounds prediction. The prediction would also have an accompanying objectiveness score at each position. The improved model demonstrated promising results at near real-time frame rates using the RPN component to inform the

unified network where to look. The introduction of RPN also helps improve the quality of region proposal and accuracy (Ren et al. 2016). Figure 5 shows the illustration of the RPN.



a)



b)

Figure 5. Region Proposal Network Illustration (a), and Example Detections using RPN Proposals (b). Source: Ren et al. (2016).

The SSD model encapsulates all computation within a single deep neural network to perform object detection using a set of preset bounding boxes of different aspect ratios for each featured map location. The model assesses and creates scores for object class detection and could modify the detection box to better fit the object. It is followed by performing non-maximum suppression to deliver the final detection results. Hongyu Liu and Bo Lang (2016) made use of "multi-scale convolutional bounding box outputs attached to multiple feature maps at the top of the network" (16). This model also does not provide proposal generation or feature resampling (Liu et al. 2016). Figure 6 presents the network architecture of SSD. The base of the network before the classification layers is built upon a standard image classification architecture. Although several networks are available, the SSD uses the VCG-16 network. Lui and Lang (2016) added extra feature structures to the network to enhance the detection capabilities.

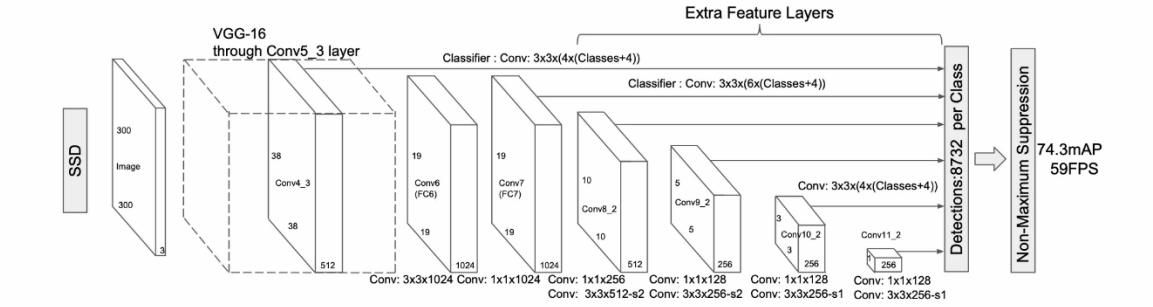


Figure 6. SSD Network Architecture. Source: Liu and Lang (2016).

A comparison of multi-sized ship detection performance among the three network models was performed by Zhong-Qui Zhao et al. (2019). It should be noted that the SSD model in their study did not utilize a professional network to provide box locating. The model composed entirely of a CNN also took convolutional features of different depths into consideration. By comparison, the YOLOv2 model in Zhao et al.'s study did not utilize RPN and had a different framework from both the R-CNN and SSD models. Their research results showed the YOLOv2 model is fast at calculation but slightly less accurate than the other models.

A separate study also demonstrated that the learning rate of YOLOv2 is 68 percent and 16 percent faster than the Faster R-CNN model and SSD model, respectively, but had poorer accuracy when used to detect small traffic signs (Grag et al. 2019).

The present research concurs with the prior studies and considers the YOLOv2 model as most suitable for real-time UXO detection. This is because the demand for a faster learning rate is paramount to support warfighters in maintaining operational momentum. Also, a high-performance graphics unit required by models processing large inputs would likely be unavailable in an operational environment.

C. SELECTION OF THE NUMBER OF ANCHOR BOXES

Both the YOLOv2 and Faster R-CNN models use multiple anchor boxes of different sizes and aspect ratios to locate an object's presence within an image. The anchor boxes estimation function obtains the Intersection over Union (IoU) distance through a k-means clustering method. IoU represents the overlap of the anchor box boundaries and the ground truth. Figure 7 shows some examples of the use of anchor boxes.

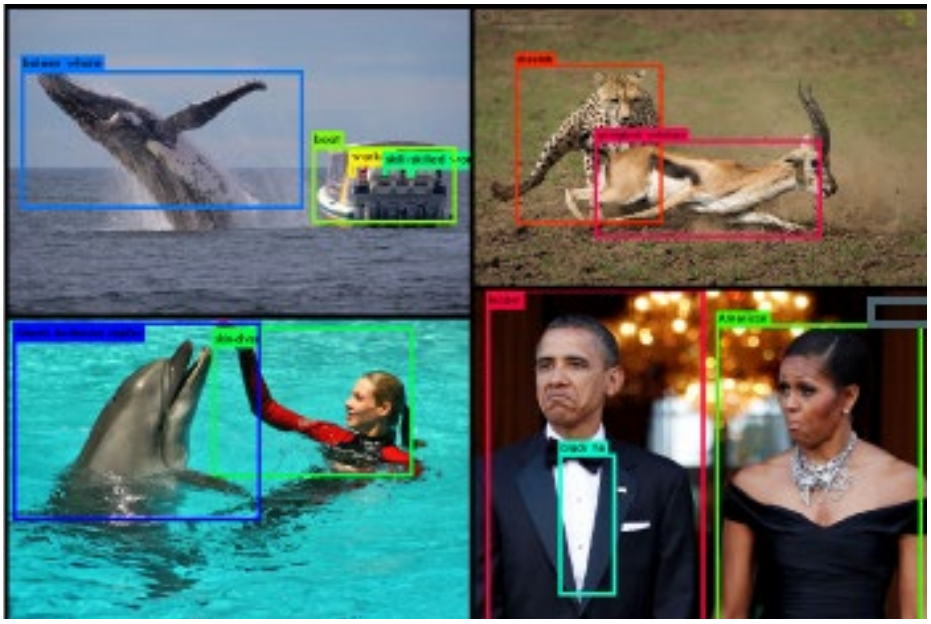


Figure 7. Examples of Anchor Boxes in the YOLOv2 Model. Adapted from Redmon and Farhadi (2016).

The quantity of anchor boxes used affects effectiveness as well as efficiency during object detection (Redmon and Farhadi 2016). The higher the number of anchor boxes, the more accurate the model will be, but it also brings about an increased period of training and might lead to degraded detection performance due to overfitting. Therefore, it is important to select a suitable quantity of anchor boxes to balance both factors.

From the results of prior research, shown in Figure 8, nine anchor boxes were deemed optimal to train the UXO detector, with a mean IoU of 0.8045, based on the relationship between the number of anchor boxes and the mean IoU (Cho 2021). The same number of anchor boxes is used in this phase of research because using more than nine anchor boxes can improve the mean IoU value trend only marginally.

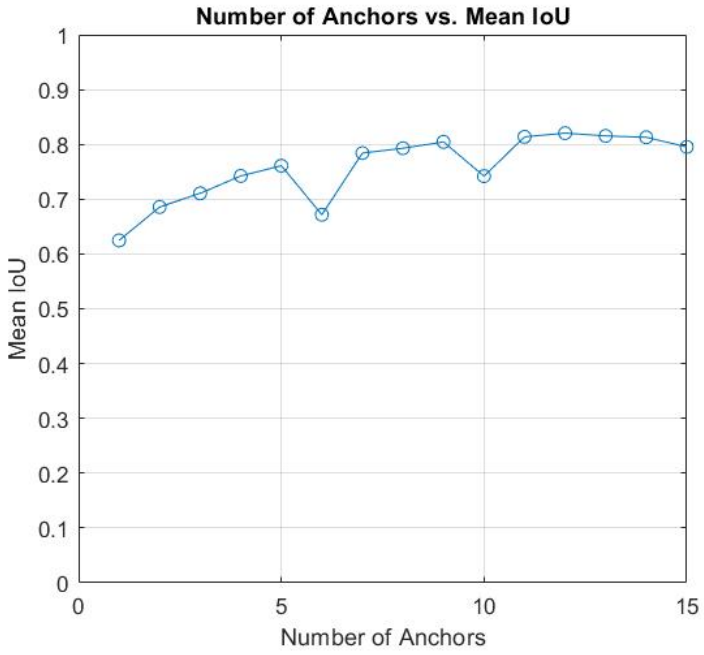


Figure 8. Number of Anchor Boxes vs. Mean IoU from the Training Data for the UXO Detector. Source: Cho (2021).

D. MEASURE OF EFFECTIVENESS

This thesis uses AP, an evaluation metric popular in object detection competitions, to measure the performance of the trained object detectors. In a blog post, Gad Ahmed

(2021) describes AP as "a method to summarize the precision-recall curve into a single value representing the average of all precisions." The metric is based on a confusion matrix to assess the performance of the detector by separating the "actual states" into different columns and "test results" into different rows (Powers 2010).

In this definition, Precision measures how accurate the predictions are, which can be obtained through the percentage of predictions that are correct

$$\text{Precision} = \frac{\text{True Detections}}{\text{All Detections}} \tag{1}$$

The Recall criterion measures how good you find all the positives

$$\text{Recall} = \frac{\text{True Detections}}{\text{Actual Occurences}} \tag{2}$$

After computing the Precision and Recall criteria, the precision-recall (PR) curve is plotted to determine the accuracy of the detector. In the PR curve, the ideal precision is 1 at all recall levels. The AP summarized the precision-recall curve where a higher AP means a better detector performance and can be computed using the 11-interpolation method as presented in the PASCAL VOC 2008 Object Detection Challenge. This means that recall values from 0 to 1 will be segmented evenly into 11 points (i.e. 0, 0.1, and 0.2). Next we compute the interpolated precision at each of the 11 points by obtaining the maximum precision measured at each recall level. Lastly, we calculate the average of the interpolated precision at each of the 11 points. (Everingham et al. 2010). The equation is as follows:

$$\text{AP} = \frac{1}{11} \sum_{\text{Re}call \in \{0, 0.1 \ldots 1.0\}} \text{Precision}_{Interpolated}(\text{Recall}) \tag{3}$$

$$\text{Precision}_{Interpolated}(\text{Recall}) = \max \text{Precision}\,(\widetilde{\text{Recall}}) \tag{4}$$

where $\max \text{Precision}\,(\widetilde{\text{Recall}})$ refers to the highest measured precision data point selected at each recall point. For example, if the precision values for first recall point "0"

are 1, 0.7, and 0.5, we will select the highest value of 1 instead of the other two values. We will continue with the same selection method for the remaining 10 recall points and take the average of the 11 recorded precision value.

E. METHODS OF TARGET CLASSIFICATION

This thesis explores whether detection capability is affected by using different methods of target classification. The author of the thesis on the previous research phase positioned the bounding boxes around the targets, with some boxes being loosely fitted and some fitted tightly (Cho 2021). He labeled the target within the image in a non-standardized manner and did not follow a particular methodology or pattern. The present thesis varies the methods of target classification. One method was to manually position bounding boxes tightly around the targets in all the images in the dataset using the Image Labeler application, which took approximately five hours to complete labeling the dataset. The other method was to position the bounding boxes loosely instead, which reduced the target classification time by nearly 60 percent, to 1.5 hours. Figure 9 provides a visual overview of the different methods of target classification applied in this research.

Nevertheless, other target classification methods are available in the Image Labeler application, for example, pixel region of interest (ROI) labels and projected cuboid labels. The pixel ROI labeling method assigns a target through pixels for semantic segmentation; users can use polygons, brushes, or flood fill to aid with their target classification. Users can also "draw" 3-D bounding boxes around their target of interest for the projected cuboid labels method. Figure 10 shows these other available methods of UXO labeling. Labels could also be broken down to sub labels to provide a greater level of detail about the ROIs in the labeled ground truth data. For example, a UXO label might contain warhead, body, and propulsion sublabels. When more details cannot be drawn, attributes can be included to specify more information about the target, like the color and type of UXO.

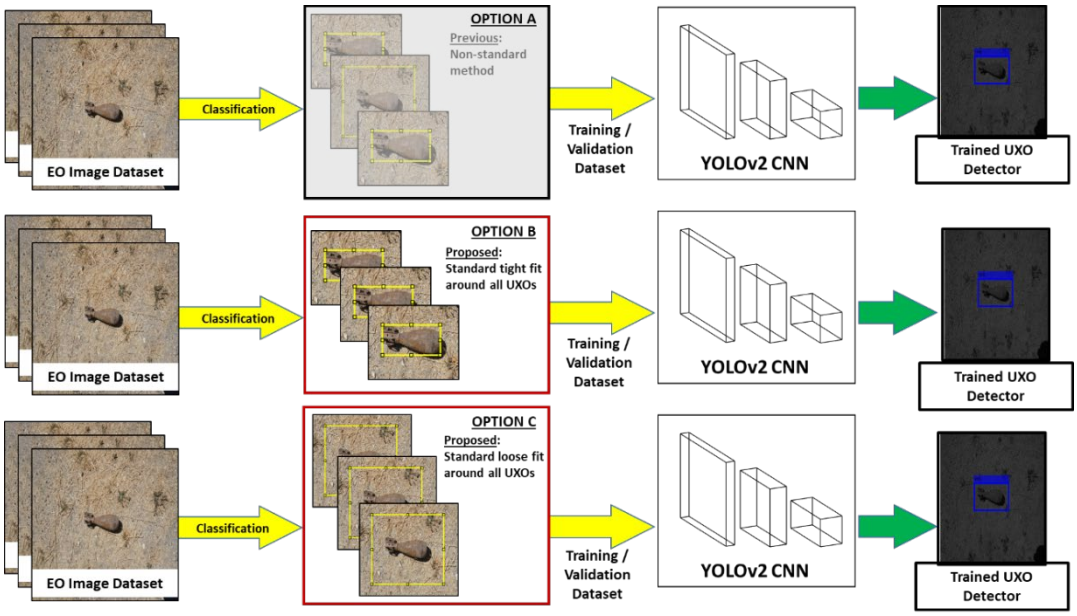
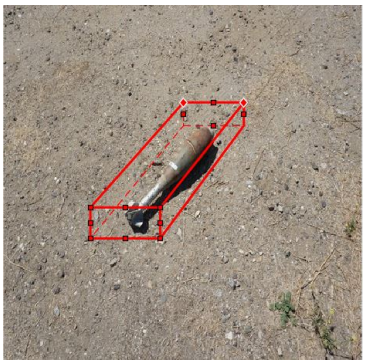


Figure 9. Different Methods of Target Classification

a)



b)

Figure 10. Alternate Methods of UXO Labeling [Projected Cuboid Method (a), and Pixel Method (b)]

THIS PAGE INTENTIONALLY LEFT BLANK

IV. DATA SOURCES AND DATA PROCESSING METHODS

This chapter describes the data sources and equipment as well as the data processing methods used for this research.

A. DATA SOURCES

This research utilized a total of 1,225 of single-spectrum RGB images obtained from the preceding phase of the research conducted at the Naval Postgraduate School by Cho (2021) to train, validate, and test the sUAS-based UXO detector model.

B. UTILIZED EQUIPMENT

This research utilized commercially available products to collect and process data.

1. Computation Platform

The MATLAB programming tool was used to perform UXO detection via system algorithm. All data processing and computations were performed on a generic laptop for the study to replicate a similar processing capacity in the area of operations. The relevant system specifications are shown in Table 1.

Table 1. Computing System Specification. Adapted from Hewlett-Packard (n.d.).

Hewlett-Packard 14	
Processor	Intel® Core™ i5-5200U (2.2GHz, 2 cores)
Memory	8 GB RAM
Graphic card	NVIDIA GeForce 820M

2. EO Imaging System

For the prior research phase on which this thesis builds, Cho used the Sony Alpha a6000, a commercial digital camera, to collect the RGB data (Cho 2021). Due to operational restrictions, the images captured by the digital camera are regarded as images obtained from the sUAS equipped with an EO sensor for the purposes of the present research. The relevant specifications of the camera are shown in Table 2.

Table 2. Relevant Specifications of Imaging System. Adapted from Sony (n.d.)

Sony Alpha a6000	
EO Spectrum	Single (Merged RGB)
Sensor Type	Complementary Metal-Oxide-Semiconductor
Max View Angle	83°
Video Resolution	1920 x 1080 (@60 fps)
Image Resolution	6000×4000 (24 Mega-pixel)
Image Shutter Speed	1/4000 to 30 sec

C. DATA PROCESSING METHODS

This section describes the data processing methods used to curate the collected images. There are four processing steps to prepare the data; first, the EO images data from the digital camera are resized according to the optimal dimensions for the selected DLCNN, and the ground truth is then labeled within the resized dataset. The labeled dataset is then allocated to training or validation and test subsets. Finally, image augmentation is performed on the training dataset.

1. Resizing the Images

The selected YOLOv2 CNN model can be trained using images of different sizes, but the image data used for this research were resized from their original dimensions to 416 x 416 pixels as illustrated in Figure 11. The YOLOv2's convolutional layers reduced spatial resolution by a factor of 32 while maintaining a two-dimensional representation of the image to produce an output feature map of 13 x 13 pixels. This in turn creates an odd quantity of locations that can accommodate a single cell in the center of the image. Objects are often located at the center of images, so it is good to have a singular center location instead of four close locations (Redmon and Farhadi 2016). This allows for better CNN performance when training with this image size and reduces the training time of the model, which can make it more efficient in the operational environment.

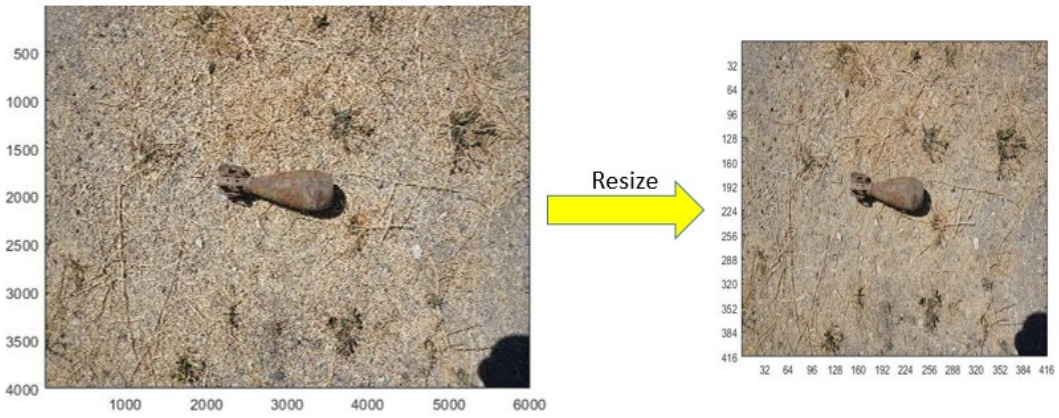


Figure 11. Original Image Resizing

2. Labeling the Ground Truth

Ground truth labeling refers to identifying intended objects and targets within the image data. This ground truth is used to train the detector for UXO detection. An object detection application, Image Labeler available in the MATLAB programming tools, was used to label the ground truth within the image data set. Figure 12 illustrates the process of ground truth labeling of UXO targets present within the EO image.

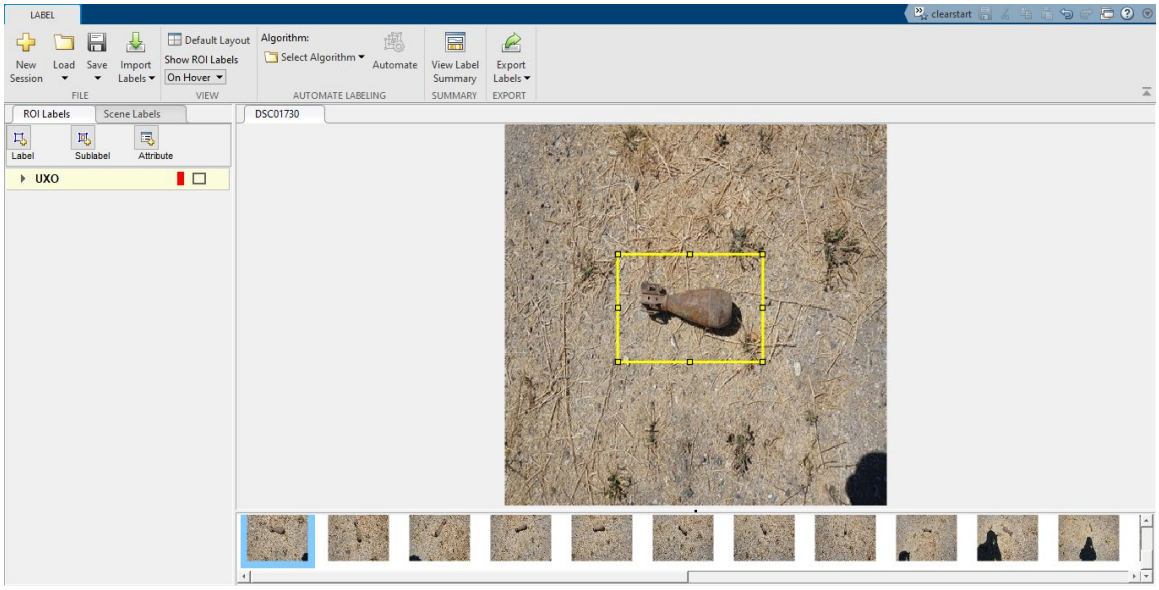


Figure 12. Ground Truth Labeling

3. Allocating the Data into Sub-Sets (Training, Validation, and Testing)

Each subset containing 1,225 single spectrum images is allocated randomly into three data sets, namely, 1) the training set, 2) the validation set, and 3) the test set. This allocation by data set is broken down by 70 percent, 15 percent, and 15 percent, respectively. The training set is meant to train the individual UXO detector, while the validation set helps to perform self-correction at regular intervals during training to improve the detector's performance. The test set is used to evaluate the trained detector.

4. Augmenting the Training Dataset

DLCNN training relies heavily on a high volume of data and large amount of resources in terms of time and budget, which are necessary to obtain a sufficient dataset (Wang et al. 2018). In real scenarios, there could be other limitations such as security consideration that prevent the collection of sufficient data for detector training. In such cases, image augmentation is a feasible option to increase data diversity through geometric manipulation like image rotation, flipping, and shifting and color adjustments in data pre-processing (Ho et al. 2019). An under-constrained model due to the lack of training data causes overfitting because of high variance, and this condition can result in an optimistic and high variance estimation of trained detector performance (Brownlee 2020). Figure 13 shows an example of the results of augmenting one EO image to present data for four EO images.

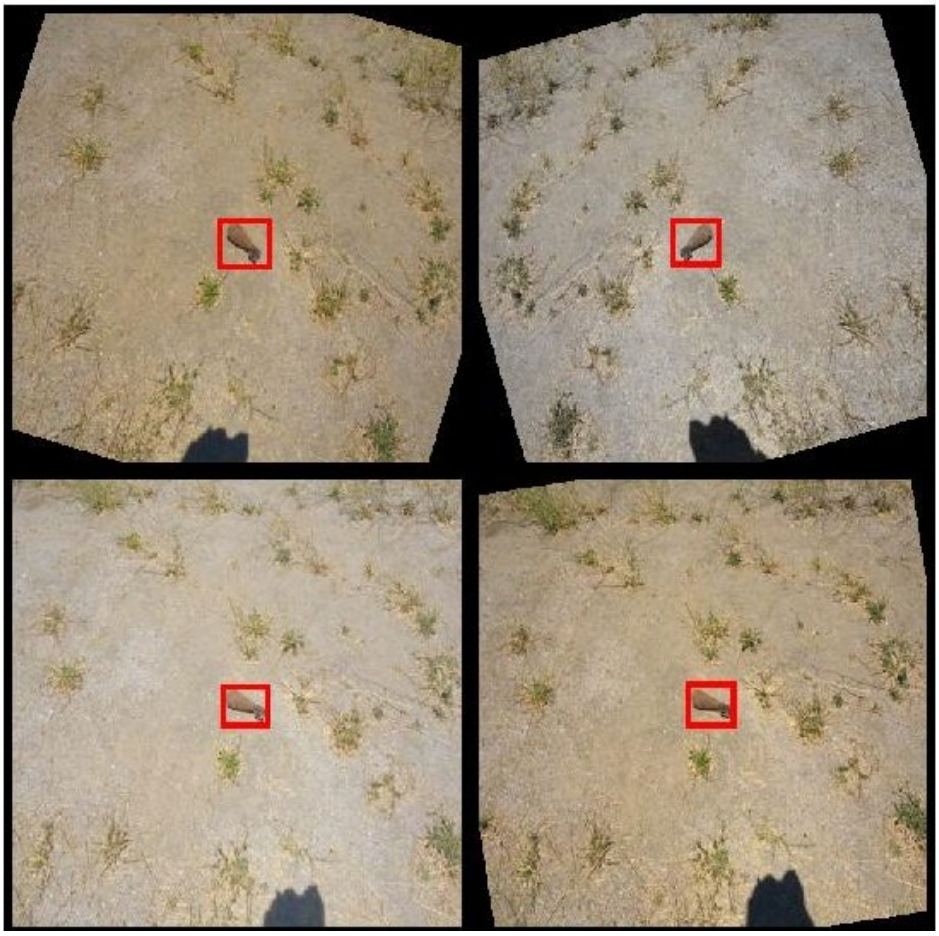


Figure 13. Example of Data Augmentation Results

THIS PAGE INTENTIONALLY LEFT BLANK

V. UXO DETECTION EXPERIMENT

This chapter describes the UXO detection experiment using the EO image dataset. EO images captured by the digital camera from the first phase of the research, as documented by Cho (2021), were 1) resized, 2) target classified, 3) divided into different data subsets, and 4) augmented. Next, the UXO detector was modeled using MATLAB and using the training and validation dataset for training of the detector. After which, the trained detector was used to detect the UXO for the test dataset. Finally, the detector's performance was evaluated according to the AP metric using an IoU threshold of 0.5, which meant the detector would eliminate any detection below the IoU of 0.5. The MATLAB codes used to train and evaluate the UXO detector using the tight-fitted labeled dataset is included in the Appendix; the codes used were the same for training and testing of the loose-fitted labeled dataset except for updates to the relevant fields from tight-fitted to loose-fitted.

A. DATA PROCESSING

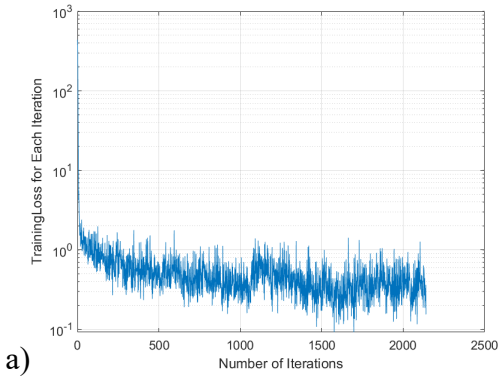
This research phase used the 1,225 UXO images in combined RGB format from the preceding phase documented by Cho (2021). These EO images were resized to the optimum image size (480 x 480 pixels) for a YOLOv2 network. Next, the ground truth, which is the UXO location in the images, was classified and labeled using the Image Labeler. There were two types of target classification in this thesis. For the first dataset class, the bounding boxes were positioned as tightly as possible around the UXO in the images. For the second dataset class, the bounding boxes were deliberately positioned to provide about two times the size of the UXO in the image. The labeled UXO datasets were then allocated randomly into the training, validation, and test subsets. Lastly, the assigned training data were rotated, skewed, and their color and contrast were adjusted to augment the training data by four times.

B. DETECTOR TRAINING

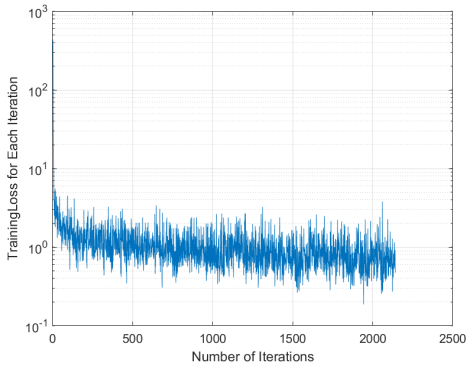
The ResNet50 and the Activation 40 Recertified Linear Unit layer were specified as the backbone network and for feature extraction, respectively, for the YOLOv2 detector.

The Adam method was selected for the optimizer, which Diederik Kingma and Jimmy Ba (2017) defined as “an algorithm for first-order gradient-based optimization of stochastic objective functions based on adaptive estimate of lower-order moments” (2). The Adam algorithm further splits the training data into mini-batches, and one epoch means that the training algorithm uses mini-batches to pass through the training data. In the preceding research phase, it was shown that insignificant training loss was achieved when the mini-batch size for the detector was specified as eight and the maximum number of epochs was selected as 20 to allow fair results comparison (Cho 2021). For this reason, the same specifications were used in the present research.

The training loss of the detector trained using tight-fitted bounding boxes and loose-fitted bounding boxes were 0.1548 and 0.4716, respectively. The training loss plots are shown in Figure 14. Both training losses were close to zero and therefore can be concluded as decent training quality.



a)



b)

Figure 14. Training Loss Plot [Tight-fitted (a), and Loose-fitted (b)]

C. DETECTION DEMONSTRATION

UXO detection was performed for the two different datasets that were labeled differently. This thesis demonstrates how the detectors work by setting the detectors to detect UXO in random images from the test data subset. The detectors generate a visual bounding box around the suspected UXO appearing in the image and provide a confidence score of the detection. As the IoU threshold was specified as 0.5, any detections with scores less than 0.5 were eliminated from the detection results. Figure 15 illustrates the demonstration of UXO detection from the test dataset along with the corresponding scores.

a)

b)

Figure 15. Demonstration of UXO Detections [Tight-fitted Data (a), and Loose-fitted Data (b)]

D. EVALUATION OF DETECTOR

Evaluation of both detectors that were trained using different datasets was performed using the MATLAB computer vision toolbox by plotting the PR curve as illustrated in Figure 16. The AP for the tight-fitted detector and loose-fitted detector were evaluated to be 0.904 and 0.895, respectively. It shows a one percent improvement in detection performance using the same dataset. This superiority of the model is based on 1) classification metric performance, which identifies whether the object is present in the image, and 2) localization metric performance, which better predicts the coordinates of the bounding box with the ground truth around the object present in the image. The steep drop in precision value at the beginning of the loose-fitted PR curve could be due to a high probability of False Positive detection return, and as the precision value is inversely proportionate to the False Positive count, it caused the precision value to drop steeply.

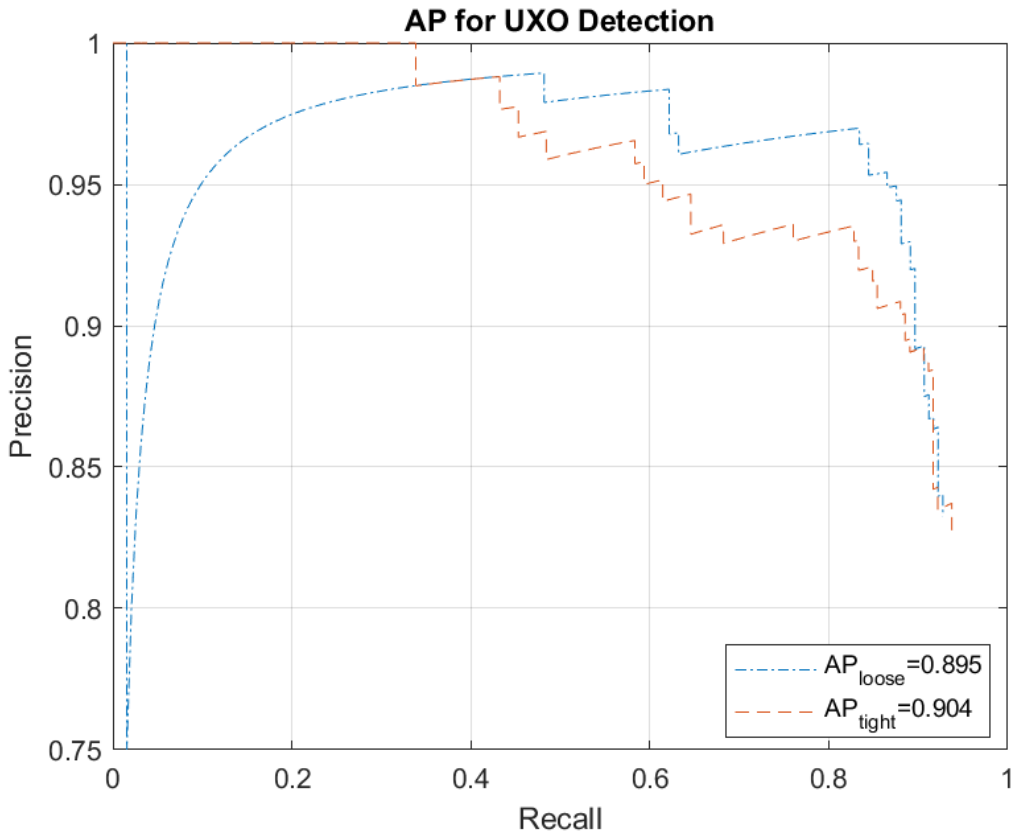


Figure 16. Combined Precision-Recall Graphs

VI. INTEGRATION OF MS SENSOR WITH SUAS

This chapter describes the integration process for a MS sensor with an sUAS. The integration offers a phased approach to implement the operational capability using a better detection capability—the MS sensor.

A. MULTI-SPECTRAL SENSOR: MICASENSE REDEDGE-MX™

The MicaSense RedEdge-MX MS sensor uses separate lenses to obtain five discrete spectral bands simultaneously: red, green, blue, near infrared, and red edge spectrum. The five spectrums capture narrow bands of blue (centered on 475 nm), green (560 nm), red (668 nm), near infrared (NIR, 840 nm), and red edge (717 nm) (MicaSense n.d.). An important feature of the sensor is the Downwelling Light Sensor (DLS) 2. The DLS 2 unit is an advanced incident light sensor that measures ambient light and sun angle in the environment to compensate the lighting index for the images to improve image quality. The DLS 2 also provides locality information of the images captured (i.e., GPS, altitude, etc.). The image and relevant specifications of the sensor are shown in Figure 17 and Table 3, respectively.

Figure 17. MicaSense RedEdge-MX™. Source: MicaSense (n.d.).

Table 3. Relevant Specifications of MicaSense RedEdge-MX™. Adapted from MicaSense (n.d.)

Spectral Bands	Blue, Green, Red, Red edge, Near-IR (narrowband)
RGB Color Output	Global Shutter, aligned with all bands
Ground Sample Distance	8 cm per pixel (per band) at 120 m (~400 ft) AGL
Capture Rate	1 capture per second (all bands), 12-bit RAW
Field of View	47.2° Horizontal and 35.4° Vertical
Resolution	1280 x 960 (1.2 MP x 5 imagers)
Weight	231.9 g (8.18 oz.) - includes associated accessories
Sensor Dimensions	8.7 cm x 5.9 cm x 4.54 cm (3.4in x 2.3in x 1.8in)

B. SMALL UNMANNED AERIAL VEHICLE: DJI INSPIRE 1 PRO

The DJI Inspire 1 is considered a Group 1 UAS (i.e., an sUAS). It is listed as a professional grade product based on the manufacturer's website and is meant for taking aerial photos or videos. Table 4 summarizes the relevant specifications of the DJI Inspire 1.

Table 4. Relevant Specifications of Inspire 1 Drone. Adapted from Adapted from DJI (n.d.)

Max Speed	49 mph or 79 kph (Attitude mode, no wind)
Max Angular Velocity	Pitch: 300°/s and Yaw: 150°/s
Max Flight Time	~18 mins
M ax Wind Speed Resistance	10 m/s
Weight	6.74 lbs (3060 g) - includes associated accessories
Max Transmission Range	3.1 miles or 5 km (unobstructed, free of interference)

C. INTEGRATION

The MicaSense MS sensor has been specifically designed to allow ready-to-use integration with popular UAS, and DJI Inspire 1 was one of the commercial sUAS specifically identified to allow straightforward integration.

The DLS 2 unit is first mounted at a position near the top of the drone because it requires an unobstructed path to measure the ambient sunlight intensity and sun's angle and record these environmental metadata obtained by MS sensor during flight. An image processing software could be used to utilize the recorded data to correct for ambient lighting changes during flight, for example during cloudy environment. Next, a sensor mounting plate is affixed to the back of the MS sensors for the quick mount adaptor to be

attached to the drone's camera port. The quick mount adaptor contains an electrical interface to provide power for the MS sensor through the drone's power source. Lastly, the DLS 2 unit is connected to the MS sensor via an electronic connector board to allow the transfer of data like GPS coordinates from the DLS 2 to the memory storage of the MS sensor.

Figure 18 shows the actual integrated system composed of the MS sensor and the DJI Inspire drone that can perform data collection. The YOLOv2 DLCNN is designed to perform real-time object detection (Redmon and Farhadi 2016). A real-time UXO detection system would improve the efficiency of the detection process but at the same time require more resources such as additional equipment and processing power. Figure 19 presents an example of an MS image captured by the integrated system at a flight altitude of 10 meters.

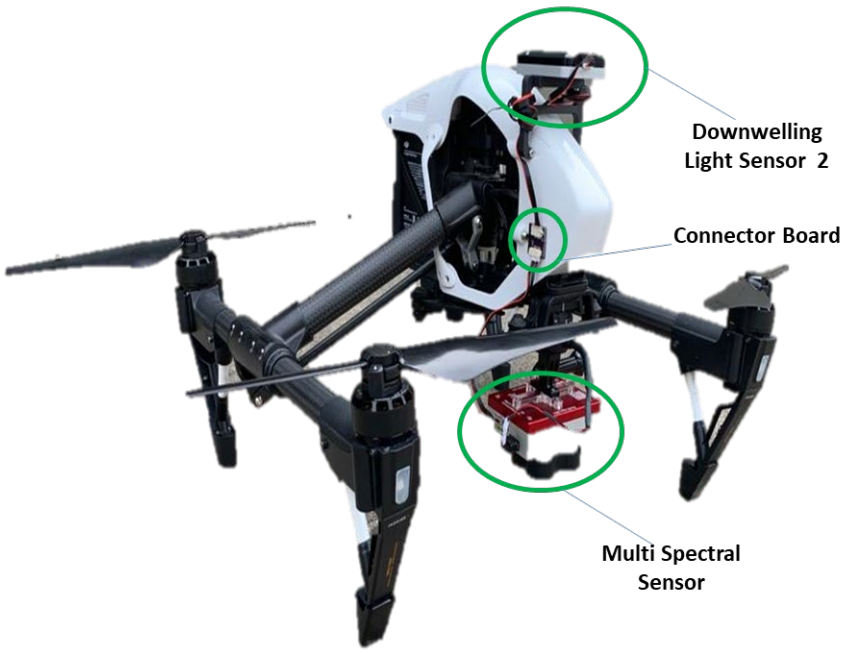


Figure 18. Actual MS Sensor Integrated with DJI Inspire Drone

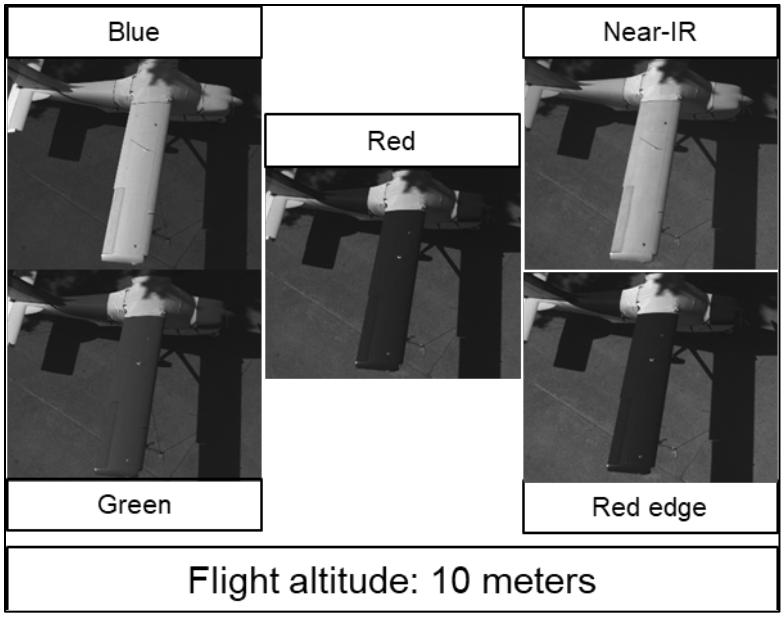


Figure 19. Example of MS Image Captured by the Integrated System

Unfortunately, the integration process and further development were disrupted by a 2020 ban imposed by the National Defense Authorization Act (NDAA) on U.S. DOD's operation of sUAS manufactured in a covered foreign country or by an entity domiciled in a covered foreign country.

VII. CONCLUSIONS AND RECOMMENDATIONS

This chapter summarizes the research findings and compares the detection performance of the different target classification methods and poses recommendations for future work that would advance the proposed UXO detection system in this thesis.

A. SUMMARY OF FINDINGS

This thesis explores the impact of different data processing factors when employing an AI-based detection algorithm for data collected by an sUAS. Specifically, it aims to understand the effects of UXO classification during the labeling of ground truth on the performance of the detector.

The research and experiments performed in this thesis revealed the following:

- A YOLOv2 network detector trained with a dataset composed of deliberately tight-fitted bounding boxes to label UXO was able to detect UXO successfully. The precision and recall evaluation of this trained detector indicated 0.904 AP against the test dataset. This could be considered relatively good performance for the object detection tasks.

- The other YOLOv2 network detector trained with a dataset composed of deliberately loose-fitted bounding boxes to label UXO was also able to detect UXO successfully. Nonetheless, its detection performance was lower than that of the previous detector, and its precision and recall evaluation indicated 0.895 AP against the test dataset.

- Using Cho's (2021) evaluation results from the prior research phase that labeled the UXO by positioning the bounding boxes around the targets, some boxes were loosely fitted and some were tightly fitted in a non-standardized manner and did not follow a particular methodology. The evaluation of the trained detector showed 0.774 AP against the test data (Cho 2021).

- A comparison of the detection performance of the three different methods of UXO labeling showed that the deliberate methods (both tight-fitted and loose-fitted) attained better detection performance than the non-standard method. This showed that the standardized method of labeling might be more beneficial for detector performance.

- The recommended method of labeling would be loose-fitted bounding boxes, because the time required for loose-fitted UXO labeling is about 60 percent of that required for tight-fitted labeling, with only a marginal drop in detection performance.

B. RECOMMENDATIONS FOR FUTURE WORK

Future research relating to this topic should consider expanding on the following areas of this study to understand field better:

- With the successful integration of the MS sensor with the sUAS, research efforts could aim to develop a real-time UXO detection system with better detection performance. It is essential, however, that the prohibition on operation or procurement of foreign-made unmanned aircraft systems be removed or amended to permit limited cost-effective research work to be completed.

- More improvements to the detection capabilities could be explored through the variation of target classification methods. At the same time, the increase in man-hours needed to perform detailed labeling might deliver only marginally superior detection capabilities. Therefore, more experiments could be performed to recommend the optimal method for target classification for UXO detection.

APPENDIX. MATLAB CODE FOR UXO DETECTOR TRAINING AND EVALUATION

The code used in the detection algorithm developed in this thesis was modified from work performed and documented by Cho at the Naval Postgraduate School in the preceding phase of the research (Cho 2021).

```
%% Load Groundtruth

load Data_UXO.mat

%% Create table data format from ground truth data
trainingDataset=Data_UXO.LabelData;
trainingDataset.files=Data_UXO.DataSource.Source;
trainingDataset=trainingDataset(:,[2,1]);

%% Visualize the training images (Labelled)

index=10;                                          %select Image # to
display

Iout=imread(trainingDataset.files{index});
for k=1:width(trainingDataset)-1
bboxes = table2array(trainingDataset(index,k+1));
if ~isempty(bboxes{1})
   Iout=insertObjectAnnotation(Iout,'rectangle',bboxes{1},...
       trainingDataset.Properties.VariableNames{k+1},'Color','red',...
       'fontsize',12,'linewidth',5);
end
end
figure, imshow(Iout);

%% Divide training data into Train/Validation/Test set
% set default for reproducible
rng(1004);
%randomize

% Divide the dataset into training set and test set
shuffledIndices=randperm(height(trainingDataset));
idx=floor(0.7*length(shuffledIndices));
trainingData=trainingDataset(shuffledIndices(1:idx),:);

validationIdx=idx+1:idx+1+floor(0.15*length(shuffledIndices));
validationData=trainingDataset(shuffledIndices(validationIdx),:);

testIdx=validationIdx(end)+1:length(shuffledIndices);
testData=trainingDataset(shuffledIndices(testIdx),:);
%% use imageDatastore and boxLabelDatastore to create datastores
```

```
imdsTrain=imageDatastore(trainingData{:,'files'});
bldsTrain=boxLabelDatastore(trainingData(:,2:end));

imdsValidation=imageDatastore(validationData{:,'files'});
bldsValidation=boxLabelDatastore(validationData(:,2:end));

imdsTest=imageDatastore(testData{:,'files'});
bldsTest=boxLabelDatastore(testData(:,2:end));

%Combine image and box label datastores
trainingData=combine(imdsTrain,bldsTrain);
validationData=combine(imdsValidation,bldsValidation);
testData=combine(imdsTest,bldsTest);

%% Define backbone network
basenetwork=resnet50();
numClasses=width(trainingDataset)-1;                          %# of class
= 1
inputSize=[416 416 3]; %3 represent 3 spectrum - might need to change
to 1

%% Anchor Box Estimation - Select Number of AnchorBoxes V.S. the Mean
IoU

anchorTraining = boxLabelDatastore(trainingDataset(:,2:end));
maxNumAnchors=15;
meanIoU=zeros([maxNumAnchors,1]);
anchorBoxes=cell(maxNumAnchors,1);

doTrain= true;
if doTrain
for k = 1:maxNumAnchors
   [anchorBoxes{k},meanIoU(k)]=estimateAnchorBoxes(anchorTraining,k);
end
save AnchorUXO anchorBoxes meanIoU
else
load AnchorUXO.mat
end
figure
plot(1:maxNumAnchors,meanIoU,'-o')
ylabel("Mean IoU")
xlabel("Number of Anchors")
title("Number of Anchors vs. Mean IoU")
grid on

%% Specify the number of anchor box
numAnchors=9;
allBoxes=round(cell2mat(reshape(table2array(trainingDataset(:,2:end)),.
..
[],1)));
scale=inputSize(1:2)./size(imread(trainingDataset.files{1}),[1 2]);
anchorBoxes = round(anchorBoxes{numAnchors}.*scale);
meanIoU_chosen = meanIoU(numAnchors);
```

```
%% Choose Network to Train            %allow network change for future
use
nettotrain = 1;                       %1 = YOLOv2, 2 = SSD, 3 =
FasterRCNN


%% Create Network (YOLOv2 or SSD or FasterRCNN)
if nettotrain ==1
featureLayer='activation_40_relu';
lgraph=yolov2Layers(inputSize,numClasses,anchorBoxes,basenetwork,...
   featureLayer);
elseif nettotrain ==2
lgraph=ssdLayers(inputSize, numClasses, 'resnet50');
else
lgraph=fasterRCNNLayers(inputSize, numClasses, anchorBoxes,...
   'resnet50');
end

%% Data Augmentation (function at the end of codes)
augmentedTrainingData=transform(trainingData,@augmentData);

augmentedData = cell(4,1);
for k = 1:4
data=read(augmentedTrainingData);
augmentedData{k}=insertShape(data{1},'Rectangle',data{2},...
   'linewidth',5,'Color','red');
reset(augmentedTrainingData);
end
figure
montage(augmentedData,'BorderSize',10,'Size',[2 2])       %for demo
purpose

%% Preprocess Training Data
preprocessedTrainingData=transform(augmentedTrainingData,...
@(data)preprocessData(data,inputSize,nettotrain));
preprocessedValidationData=transform(validationData,...
@(data)preprocessData(data,inputSize,nettotrain));

%% Configure the network training options      %specify detector
parameters
options = trainingOptions('adam',...
     'InitialLearnRate',0.001,...
     'Verbose',true,...
     'MiniBatchSize',2,...
     'MaxEpochs',2,...                 %Single Spectrum - use 20
     'Shuffle','never',...
     'VerboseFrequency',10,...
     'ValidationData',preprocessedValidationData);

%% Train the network                                  %options for future
use

doTrain=false; %false
if doTrain
```

```
if nettotrain == 1

[detector,info]=trainYOLOv2ObjectDetector(preprocessedTrainingData,lgra
ph,options);
   save UXO_yolo_detector
elseif nettotrain ==2

[detector,info]=trainSSDObjectDetector(preprocessedTrainingData,lgraph,
options);
   save UXO_SSD_detector
else
   [detector,info] =
trainFasterRCNNObjectDetector(preprocessedTrainingData,lgraph,
options);
   save UXO_FasterRCNN_detector
end
detector;
figure
plot(info.TrainingLoss)
grid on
xlabel('Number of Iterations')
ylabel('TrainingLoss for Each Iteration')
else
load UXO_yolo_detector.mat             %to load the trained detector
end

%% Read a test image into the workspace
I=read(testData);
I=imresize(I{1},1);                  %%use first picture of testdata
[bboxes,scores,labels]=detect(detector,I,'threshold',0.6);
                                        %thres=0.4 when combineMS

%% Display the results.
[~,ind]=ismember(labels,detector.ClassNames);
if(~isempty(bboxes))
I = insertObjectAnnotation(I,'rectangle', bboxes,...
   strcat(string(labels), " : ",...
   string(scores)), "LineWidth," 5, "fontsize," 12, 'Color','red');
end
figure
clf
imshow(I)

%% Evaluate Detector Using Test Set
preprocessedTestData =
transform(testData,@(data)preprocessData(data,...
inputSize,nettotrain));

%Run the detector on all the test images
detectionResults = detect(detector, preprocessedTestData,
'threshold',...
0.5,'ExecutionEnvironment','cpu');

%Evaluate the object detector using AP metric.
```

```
[ap, recall, precision] =
evaluateDetectionPrecision(detectionResults,...
preprocessedTestData, 0.5);

%Plot the PR curve.
figure
plot(recall, precision)
xlabel('Recall')
ylabel('Precision')
grid on
temp = sprintf(' = %.3f', ap);
title(sprintf(['AP for ',
trainingDataset.Properties.VariableNames{2},...
temp]));

%% my functions

%func for preprocess
function data = preprocessData(data,targetSize,nettotrain)

% Resize image and bounding boxes to the targetSize(not needed for
YOLOv2)
if nettotrain==2
scale=targetSize(1:2)./size(data{1},[1 2]);
data{1}=imresize(data{1},targetSize(1:2));
data{2}=bboxresize(data{2},scale);
else
data{1}=imresize(data{1},1);
data{2}=bboxresize(data{2},1);
end
end

function B = augmentData(A)                    %func for augmentation
B=cell(size(A));

I=A{1};
sz=size(I);
if numel(sz)==3 && sz(3) ==3
I=jitterColorHSV(I,...
    'Contrast',0.2,...
    'Hue',0,...
    'Saturation',0.1,...
    'Brightness',0.2);
end

tform = randomAffine2d('XReflection',true,'Scale',[1 1.2],
'Rotation',...
[-20 20]);
rout=affineOutputView(sz,tform,'BoundsStyle','CenterOutput');
B{1}=imwarp(I,tform,'OutputView',rout);

[B{2},indices]=bboxwarp(A{2},tform,rout,'OverlapThreshold',0.25);
B{3}=A{3}(indices);
```

```
if isempty(indices)
B=A;
end
end
```

LIST OF REFERENCES

Ahmed, Gad. 2021. "Evaluating Object Detection Models Using Mean Average Precision." *KDnuggets* (blog), March 2021. https://www.kdnuggets.com/2021/03/evaluating-object-detection-models-using-mean-average-precision.html.

Australia Department of Defense. n.d. "Types of Unexploded Ordnance." Accessed May 22, 2021. https://www.defence.gov.au/UXO/What/Types.asp.

Bekmezci, İlker, Eren Şentürk, and Tolgahan Türker. 2016. "Security Issues in Flying Ad-Hoc Networks (FANETs)." *Journal of Aeronautics and Space Technologies* 9 (2): 13–21.

Berni, Jose A. J., Pablo J. Zarco-Tejada, Lola Suarez, and Elias Fereres. 2009. "Thermal and Narrowband Multispectral Remote Sensing for Vegetation Monitoring from an Unmanned Aerial Vehicle." *IEEE Transactions on Geoscience and Remote Sensing* 47 (3): 722–38. https://doi.org/10.1109/TGRS.2008.2010457.

Bertrand, Harold E., David C. Heberlein, John T. Frasier, and Sherryl S. Zounes. 2004. "Report of UXO Technology Subgroup: Overview and Technology Assessment," IDA Document D-3007. Alexandria, VA: Institute for Defense Analyses.

Bond, Raymond, Ansgar Koene, Alan Dix, Jennifer Boger, Maurice D. Mulvenna, Mykola Galushka, Bethany Waterhouse Bradley, Fiona Browne, Hui Wang, and Alexander Wong. 2019. "Democratization of Usable Machine Learning in Computer Vision." http://arxiv.org/abs/1902.06804.

Brownlee, Jason. 2019. "A Gentle Introduction to Object Recognition with Deep Learning." *Machine Learning Mastery* (blog). May 21, 2019. https://machinelearningmastery.com/object-recognition-with-deep-learning/.

———. 2020. "Impact of Dataset Size on Deep Learning Model Skill and Performance Estimates." *Machine Learning Mastery* (blog). August 25, 2020. https://machinelearningmastery.com/impact-of-dataset-size-on-deep-learning-model-skill-and-performance-estimates/.

Carter Center, The. 2020. "Explosive Weapons Use in Syria, Report 2." January 2020. https://www.cartercenter.org/resources/pdfs/peace/conflict_resolution/syria-conflict/explosive_weapons_contamination_syria_report2.pdf.

Curran, Paul J., Jennifer L. Dungan, and Henry L. Gholz. 1990. "Exploring the Relationship between Reflectance Red Edge and Chlorophyll Content in Slash Pine." Tree Physiology 7 (1-2-3-4): 33–48. https://doi.org/10.1093/treephys/7.1-2-3-4.33.

Cho, Seungwan. 2021 "AI-Based UXO Detection Using sUAS Equipped with a Single- or Multi-Spectrum EO Sensor." Master's thesis, Naval Postgraduate School. https://calhoun.nps.edu/handle/10945/67116.

deSmet, Timothy, Alex Nikulin, William Frazaer, Jasper Baur, Jacob Abramowitz, Daniel Finan, Sean Denara, Nicholas Aglietti, and Gabriel Campos. 2018. "Drones and 'Butterflies': A Low-Cost UAV System for Rapid Detection and Identification of Unconventional Minefields." *Journal of Conventional Weapons Destruction* 22 (3) (Article 10): 1–9. https://commons.lib.jmu.edu/cisr-journal/vol22/iss3/10.

DJI. n.d. "Inspire 1 Specifications." Accessed July 14, 2021. https://www.dji.com/inspire-1.

Dorafshan, Sattar, Robert Thomas, Calvin Coopmans, and Marc Maguire. 2018. "Deep Learning Neural Networks for SUAS-Assisted Structural Inspections: Feasibility and Application." In *IEEE International Conference on Unmanned Aircraft Systems*. https://doi.org/10.1109/ICUAS.2018.8453409.

Esposito, Floriana, and Donato Malerba. 2001. "Machine Learning in Computer Vision," *Applied Artificial Intelligence* 15 (8) (November): 693–705. https://doi.org/10.1080/088395101317018546.

Etter, Delores, and Bill Delaney. 2003. "Report of the Defense Science Board Task Force on Unexploded Ordnance." Defense Science Board Task Force. Washington, D.C.: Office of the Under Secretary of Defense (Acquisition, Technology & Logistics). https://dsb.cto.mil/reports/2000s/ADA419970.pdf.

Everingham, Mark, Luc Van Gool, C. K. I. Williams, J. Winn, and Andrew Zisserman. 2010. "The Pascal Visual Object Classes (VOC) Challenge." *International Journal of Computer Vision* 88 (2) (September): 303–38. https://doi.org/10.1007/s11263-009-0275-4.

Gani, Md Osman, Somenath Kuiry, Alaka Das, Mita Nasipuri, and Nibaran Das. 2021. "Multispectral Object Detection with Deep Learning." In *Computational Intelligence in Communications and Business Analytics*, edited by Paramartha Dutta, Jyotsna K. Mandal, and Somnath Mukhopadhyay, 1406:105–17. https://doi.org/10.1007/978-3-030-75529-4_9.

Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. 2014. "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation." In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 580–87. https://doi.org/10.1109/CVPR.2014.81.

Grag, Priya, Debapriyo Roy Chowdhury, and Vidya N. More. 2019. "Traffic Sign Recognition and Classification Using YOLOv2, Faster RCNN and SSD." In *10th International Conference on Computing, Communication and Networking Technologies* (ICCCNT), 1–5. https://doi.org/10.1109/ICCCNT45670.2019.8944491.

Grone, Philip. 2003. "Definitions Related to Munitions Responsive Actions." Official memorandum. Washington, D.C.: Department of Defense. https://www.epa.gov/sites/production/files/documents/mrp_definitions_12-18-03.pdf.

Hewlett-Packard. n.d. "Laptop Computers." Accessed July 11, 2021. https://www.hp.com/us-en/shop/cat/laptops.

Ho, Daniel, Eric Liang, Ion Stoica, Pieter Abbeel, and Xi Chen. 2019. "Population Based Augmentation: Efficient Learning of Augmentation Policy Schedules." http://arxiv.org/abs/1905.05393.

Ioffe, Sergey, and Christian Szegedy. 2015. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift." http://arxiv.org/abs/1502.03167.

Khan, Asharul Islam, and Salim Al-Habsi. 2020. "Machine Learning in Computer Vision." *Procedia Computer Science* 167: 1444–51. https://doi.org/10.1016/j.procs.2020.03.355.

Kingma, Diederik P., and Jimmy Ba. 2017. "Adam: A Method for Stochastic Optimization." *2015 International Conference on Innovative Research*. http://arxiv.org/abs/1412.6980.

Lee, Wee Leong. 2018. "Assessment of Foreign Object Debris Management Using Group 1 Unmanned Aerial Systems." Master's thesis, Naval Postgraduate School. https://calhoun.nps.edu/handle/10945/60426.

Liu, Hongyu, and Bo Lang. 2019. "Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey." *Applied Sciences* 9 (20) (October): 4396. https://doi.org/10.3390/app9204396.

Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. 2016. "SSD: Single Shot MultiBox Detector." In *European Conference on Computer Vision,* edited by Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, 9905:21–37. https://doi.org/10.1007/978-3-319-46448-0_2.

MicaSense. n.d. "RedEdge – MX™." Accessed July 14, 2021. https://micasense.com/rededge-mx/.

Overton, Iain. 2020. “UK Won’t Say How Many RAF Bombs Fail to Explode - Harms Manufacturer’s ‘Trade Secrets’.” *AOAV* (blog). February 7, 2020. https://aoav.org.uk/2020/uk-government-refuses-to-release-data-on-how-many-missiles-fail/.

Ozturk, Ozan, Batuhan Sariturk, and Dursun Zafer Seker. 2020. “Comparison of Fully Convolutional Networks (FCN) and U-Net for Road Segmentation from High Resolution Imageries.” *International Journal of Environment and Geoinformatics* (IJEGEO) 7 (3) (June): 272–279. https://doi.org/10.30897/ijegeo.737993.

Powers, David. 2010. “Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation.” *International Journal of Machine Learning Technology* 2 (1): 37–63. https://arxiv.org/ftp/arxiv/papers/2010/2010.16061.pdf.

Ptak, Bartosz, and Mateusz Piechocki. 2020. “Small Object Detection and Recognition in Aerial Images.” In *GHOST Day: Applied Machine Learning Conference*. https://doi.org/10.13140/RG.2.2.29239.04007.

Qi, Zhipeng, Xiu Li, He Li, and Wentao Liu. 2020. “First Results from Drone-Based Transient Electromagnetic Survey to Map and Detect Unexploded Ordnance.” *IEEE Geoscience and Remote Sensing Letters* 17 (12): 2055–59. https://doi.org/10.1109/LGRS.2019.2962754.

Redmon, Joseph, Ali Farhadi. 2016. “YOLO9000: Better, Faster, Stronger.” In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (December 25, 2016): 6517–6525. https://www.doi.org/10.1109/CVPR.2017.690.

Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. 2016. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.” http://arxiv.org/abs/1506.01497.

Savell, Stephanie. 2019. “The Imperial Debris of War.” *Lobe Log* (blog). September 28, 2019. https://lobelog.com/the-imperial-debris-of-war/.

Sharma, Pulkit. 2019. “Image Classification vs Object Detection vs Image Segmentation.” *Medium* (blog). August 21, 2019. https://medium.com/analytics-vidhya/image-classification-vs-object-detection-vs-image-segmentation-f36db85fe81.

Sinha, Rajat, Ruchi Pandey, and Rohan Pattnaik. 2017. “Deep Learning for Computer Vision Tasks: A Review.” In *2017 International Conference on Intelligent Computing and Control (I2C2)*. https://arxiv.org/ftp/arxiv/papers/1804/1804.03928.pdf.

Sony. n.d. “Sony A6000 E-Mount Camera with APS-C Sensor.” Accessed July 11, 2021. https://www.sony.com/electronics/interchangeable-lens-cameras/ilce-6000-body-kit.

Takumi, Karasawa, Kohei Watanabe, Qishen Ha, Antonio Tejero-De-Pablos, Yoshitaka Ushiku, and Tatsuya Harada. 2017. “Multispectral Object Detection for Autonomous Vehicles.” In *Proceedings of the on Thematic Workshops of ACM Multimedia 2017 - Thematic Workshops ‘17*, 35–43. https://doi.org/10.1145/3126686.3126727.

U.S. Department of Defense. 2010 “U.S. Army Roadmap for Unmanned Aircraft System 2010–2035.” Fort Rucker, AL: U.S. Army UAS Center of Excellence. https://fas.org/irp/program/collect/uas-army.pdf.

———. 2018. “Unmanned Systems Integrated Roadmap 2017–2042.” USNI News. https://news.usni.org/2018/08/30/pentagon-unmanned-systems-integrated-roadmap-2017-2042.

U.S. Department of Transportation. 2013. “Unmanned Aircraft System (UAS) Service Demand 2015–2035: Literature Review & Projections of Future Usage.” U.S. Department of Transportation John A. Volpe National Transportation Systems Center. https://fas.org/irp/program/collect/service.pdf.

Viswanathan, Vibhuthi. 2018. “10 Latest Technology used in Artificial Intelligence.” *SpringPeople* (blog), September 5, 2018. https://www.springpeople.com/blog/10-latest-technology-used-in-artificial-intelligence/.

Wang, Tianyang, Jun Huan, and Bo Li. 2018. “Data Dropout: Optimizing Training Data for Convolutional Neural Networks.” http://arxiv.org/abs/1809.00193.

White, Eric. 2021. “Use of sUAS Technology and Imaging Sensors for Optimization of Target Weed Control in Vegetation Management Scenarios.” Master’s thesis, North Carolina State University. https://www.lib.ncsu.edu/resolver/1840.20/38704.

Wikimedia. 2007. “File:Biskeri- Camping I IMG 7238.jpg.” Accessed July 03, 2021. https://commons.wikimedia.org/wiki/File:Biskeri-_Camping_I_IMG_7238.jpg.

Zhao, Zhong-Qiu, Peng Zheng, Shou-tao Xu, and Xindong Wu. 2019. “Object Detection with Deep Learning: A Review.” *IEEE Transactions on Neural Networks and Learning Systems* 30. http://arxiv.org/abs/1807.05511.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California