# ANALYSIS OF GEOBIA ALGORITHMS FOR CONTEXTUAL DETECTION OF DPRK MISSILE TESTING FACILITIES

Ву

Connor Alec Plensdorf

A THESIS

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

Geography-Master of Science

#### ABSTRACT

# ANALYSIS OF GEOBIA ALGORITHMS FOR CONTEXTUAL DETECTION OF DPRK MISSILE TESTING FACILITIES

By

#### Connor Alec Plensdorf

Remote sensing provides people with an alternative, otherwise unattainable view to analyze the earth. Military and intelligence analysts quickly adopted this technology for tactical and strategic applications. Accordingly, these interpreters require increasingly immediate, accurate image analysis for decision-making in today's dynamic military environment. Geographic Object-Based Image Analysis (GEOBIA) provides a means for automated image interpretation modeled after the expert interpretation processes. Although the system's flexibility is advantageous for creating comprehensive image classifications, its flexibility may also preclude full automation and replication.

The goal of this research was to improve image classification outcomes in the context of missile site detection. Here a GEOBIA workflow was developed that incorporates expert human knowledge for the detection of DPRK missile testing facilities. After conducting the analyses, I determined the best-fitting parameters from those tested include the rule-based classification for the Sohae testing facility and random forest classification for Yongbyon, with no conclusive results in favor of either software. The results indicate expert human knowledge does not necessarily improve classification accuracy for this case of study sites.

Key Words: GEOBIA, image classification, contextual analysis, situation awareness, DPRK

This thesis is dedicated to the family and friends who have wholeheartedly supported me throughout this project and in past, present, and future endeavors.

## ACKNOWLEDGEMENTS

I want to thank my graduate advisor, Dr. Raechel White, for guiding me along this research project. I would also like to thank the remainder of my graduate committee, including Dr. Kyle Evered and Dr. Ashton Shortridge, for their support and effort towards this project. I also extend further gratitude to the Nuclear Threat Initiative researchers and the 38 North agency and its associated image interpreters.

# TABLE OF CONTENTS

LIST OF TABLESvii				
LIST OF FIGURESviii				
1. In	ntroduction	1		
<u>э</u> р	ackaround	0		
<b>2. D</b>	Remote Sensing for Strategic Operations	0		
2.1.	The DPRK Missile Program	С		
2.2.	The DPRK Missile Development Monitoring	6		
2.5.	Rule-Based Image Object Detection	9		
2. <del>4</del> . 2.5	Knowledge Incorporation into GEOBIA Applications			
2.5.	Feature Extraction	. <u>т</u> А		
2.0.	Objectivity			
2.7.	Workflow Reusability	.5 26		
2.0. 2 Q	Contributions of the Present Research	.0		
2.5.		. /		
2 1/	1ethods 2	a		
3. 1	Study Sitos	20		
3.1.	Data	23		
3.2.	Analysis 1 - Extraction of Expert Interpretation Cues	3		
3.4	Analysis 2 - Classification Methods Best for Missile Facility Extraction	37		
3.4	4.1 Rule-Based Classification in eCognition	9		
3 3./	1.2 Nearest Neighbor Classification in eCognition	1		
3	4.3 Random Forest Classification in eCognition	ι <u>΄</u>		
3	Analysis 3 – Comparison of Classification Software	14		
3.5.	Analysis 5 Comparison of Spatial Resolutions	17		
3.0.	Post-classification Accuracy Assessments	., .1		
5.7.		-		
4. R	esults	54		
4.1.	Knowledge Incorporation Comparison5	4		
4.1	1.1. Rule-Based Classification without Knowledge in eCognition	5		
4.1	1.2. Rule-Based Classification with Knowledge in eCognition	57		
4.1	1.3. Nearest Neighbor Classification without Knowledge in eCognition	8		
4.1	1.4. Nearest Neighbor Classification with Knowledge in eCognition	0		
4.1	1.5. Random Forest Classification without Knowledge in eCognition	51		
4.1	1.6. Random Forest Classification with Knowledge in eCognition	53		
4.2.	Software Comparison6	5		
4.2	2.1. Random Forest Classification in eCognition	5		

4.2.	2. Random Forest Classification in R	65			
4.3.	Spatial Resolution Comparison	69			
4.3.	1. Rule-Based Classification 1m	69			
4.3.	2. Rule-Based Classification 3m	71			
4.3.	3. Nearest Neighbor Classification 1m	72			
4.3.	4. Nearest Neighbor Classification 3m	73			
4.3.	5. Random Forest Classification 1m	74			
4.3.	6. Random Forest Classification 3m	75			
4.3.	7. Overall Results	77			
5. Dis	cussion				
5.1.	Limitations				
5.2.	Delimitations				
5.3.	Analysis 1 – Knowledge Incorporation Comparison	79			
5.4. Analysis 2 – Software Comparison		82			
5.5.	Analysis 3 – Spatial Resolution Comparison	83			
5.6.	Developments on Present Research	86			
6. Co	nclusion	88			
APPEND	ICES	91			
APPEN	IDIX A: Abbreviations				
APPENDIX B: Sohae Ruleset9					
APPEN	APPENDIX C: Yongbyon Ruleset94				
APPEN	NDIX D: Palisades Nuclear Energy Facility Ruleset				
REFEREN	VCES	96			

# LIST OF TABLES

Table 1: Objectives of the Study	6
Table 2: Concordance subset for the knowledge base	37
Table 3: Classifications used for analysis per feature space	39
Table 4: Features that were used for knowledge incorporation	40
Table 5: Number of samples per class for each DPRK image	42
Table 6: Knowledge Incorporation comparison accuracy assessment (above)	.64
Table 7: Knowledge Incorporation Analysis Processing Times	64
Table 8: Model 1 Accuracy Assessment	65
Table 9: Model 2 Accuracy Assessment	66
Table 10: Model 2 variables with greater than 2 significance	67
Table 11: Model 3 Accuracy Assessment	.68
Table 12: Rule-based, 1m Accuracy Assessment	70
Table 13: Rule-based, 3m Accuracy Assessment	71
Table 14: Nearest Neighbor 1m, Accuracy Assessment	72
Table 15: Nearest Neighbor 3m, Accuracy Assessment	73
Table 16: Random Forest 1m, Accuracy Assessment	75
Table 17: Random Forest, 3m Accuracy Assessment	76
Table 18: Spatial Resolution Analysis Classification Processing Times	77

# LIST OF FIGURES

Figure 1: DPRK Missile Ranges. Source: https://www.dw.com/en/which-us-cities-could-north- koreas-ballistic-missile-hit/a-39881831	- 14
Figure 2: Number of expert interpretations from 38 North per missile testing site in this study	.32
Figure 3: Sohae Image. Source: Planet Labs	.33
Figure 4: Yongbyon Image. Source: Planet Labs	.34
Figure 5: Palisades nuclear power plant, 1m resolution and 3m resolution (right). Source: USGS	.34
Figure 6: Sohae samples used for supervised and machine learning classifications	.42
Figure 7: Yongbyon samples used for supervised and machine learning classifications	43
Figure 8: Nuclear Power Plant (top) and airfield (bottom).1m images on left and 3m images on right	n 48
Figure 9: Samples used for the spatial resolution comparison classifications	51
Figure 10: Rule-based no knowledge classification, Sohae	.55
Figure 11: Rule-based no knowledge classification, Yongbyon	.56
Figure 12: Rule-based with knowledge classification, Sohae	.57
Figure 13: Rule-based with knowledge classification, Yongbyon	.58
Figure 14: NN no knowledge classification, Sohae	.59
Figure 15: NN no knowledge classification, Yongbyon	.59
Figure 16: NN with knowledge classification, Sohae	.60
Figure 17: NN with knowledge classification, Yongbyon	.61
Figure 18: RF no knowledge classification, Sohae	62
Figure 19: RF no knowledge classification, Yongbyon	62
Figure 20: RF with knowledge classification, Sohae	.63
Figure 21: RF with knowledge classification, Yongbyon	.64
Figure 22: Prediction using Model 1 (z-value of pixel)	.66
Figure 23: Prediction using RF Model 2 (RFsp with z)	.67

Figure 24: Prediction using RF Model 3 (RFsp, no z)	68
Figure 25: Rule-based classification, 1m	69
Figure 26: Rule-based classification, 3m	71
Figure 27: NN classification, 1m	72
Figure 28: NN classification, 3m	73
Figure 29: RF classification, 1m	75
Figure 30: RF classification, 3m	76

## 1. Introduction

Despite its benefits of timely and comprehensive visualization of the earth's surface, satellite imagery may not satisfy consumer needs entirely. Military and intelligence consumers often require near real-time analysis for proper tactical or strategic decision-making (Cloud and Clarke 1999). Military image analysts undertake tedious, time-consuming processes of manually analyzing images. As late as 1997, public image analysts manually analyzed imagery to detect and confirm the location of the 1974 Indian nuclear weapons test (Gupta and Pabian 1997). The 1990s declassification of CORONA satellite imagery opened up dialogues between military and public analysts, allowing greater information sharing between the domains (Cloud and Clarke 1999).

The military has used remote sensing technologies for observation, and intelligence gathering for over 100 years. During the American Civil War and the 1849 bombardment of Venice, Italy, unmanned hot air balloons conducted aerial reconnaissance for military operations (Watts, Kobziar, and Percival 2009). By World War II, the world's militaries used aerial photography, for situation awareness and damage assessment. Greater advances in military technology and increasing demands, such as the need for increased visibility and range, led to the development of more modern technologies for remote sensing (Perkins and Dodge 2009).

The Cold War drove developments in remote sensing, particularly satellite imagery, which has since dominated national security military applications. The CORONA satellite imagery program – the first US spy satellite imagery program - provided strategic means for decisionmaking against the Soviet Union throughout the Cold War (Cloud and Clarke 1999).

With satellite imagery technologies initiated in the CORONA program, the military since conducted several operations concerning developments in imagery gathering. Much of the

military use today involves the detection of hazardous entities and other forms of intelligence gathering. For instance, military personnel have used hyperspectral imagery to detect vehicles under vegetation canopies (Shippert n.d.). The use of satellites is also used to assess military success and progress. Examples of such analytical goals include the 2007 surge, in which the U.S. military deployed over 30,000 military personnel to Baghdad, Iraq, and assessments of its effectiveness in swaying the tide of the political and social battle of nation-building and reconstruction. Accordingly, the military used satellite imagery from the Defense Meteorological Satellite Program (DMSP) to detect changes in city lights visible from space before, during, and after the surge. Researchers indicated that the presence of visible light, would display a likely increase in infrastructure whereas darkness, or an absence of visible light would likely indicate a decrease in infrastructure or absence altogether (Agnew et al. 2008).

Details about the methods of image analysis used by the military remain classified. The increase in availability of commercial satellite imagery, as well as commercial initiatives for declassified imagery, such as John Pike's *Public Eye*, have improved public awareness about military remote sensing. The *Public Eye* stems from Globalsecurity.org and relates to issues of national intelligence and security and uses declassified CORONA imagery and aerial U2 imagery (Perkins and Dodge 2009). Widespread use of remote sensing imagery by the media has meant there is an increasing demand for open source imagery by the public. With programs such as *Public Eye* and Digital Globe, a commercial producer of Quickbird imagery, high-resolution satellite imagery is becoming increasingly open and public.

Geographic object-based image analysis (GEOBIA) use has risen due to its capability to improve analytical output in the face of increasing image analysis demands. GEOBIA challenges

and builds upon the largely standardized pixel-based methods, integrating computer vision and pattern recognition processes with traditional earth observation workflows (Blaschke et al. 2014). Its emergence likely stems from recent historical events, such as the 1990s changes in U.S. space policy (Hitchings 2003) and the development of more powerful computers and processing technologies (Hay and Castilla 2008).

GEOBIA is an image analysis method centering on the patterns created by pixels rather than the pixels themselves. GEOBIA began to constitute a potential paradigm shift in the early 2000s (Blaschke et al. 2014). Patterns created by these pixels are designated as image objects, which are items of interest in the image (Blaschke et al. 2014). Focusing on image objects, GEOBIA mimics the way humans interpret images, instead of focusing on individual pixels often unseen with the naked eye (Hay and Castilla 2008). The issue with this is ultimately that pixels are not features of an image and are, thereby, subject to internal homogeneity or heterogeneity, in which they capture only parts of features or located on feature edges, respectfully. With a grouping via segmentation, pixels of similar spectral value may be logically grouped into recognizable features. Hay and Castilla define GEOBIA as a sub-discipline of geographic information science (GIScience) that automates segmentation of images and evaluates the spatial and spectral characteristics of the image objects created. Results may be used in a GIS-ready format (Hay and Castilla 2008).

GEOBIA begins by dividing entire images into candidate image objects through segmentation. Segmentation is the process of dividing the image into image-objects of spectrally-similar pixel groups. Classification, on the other hand, is the process of assigning these image-objects classes to represent a given analysis. Pre-built algorithms, such as the multiresolution segmentation and chessboard segmentation, and are used for this segmentation, after which the analyst creates objects from segmented images with trial-and-error rule building (Belgiu, Hofer and Hofmann 2014). During GEOBIA's second phase, the image is classified into user-designated classes based on classification rules. This step permits the user to draw from his/her knowledge of the landscape and subject matter to create threshold-based rules to assign specific classes across the image (Belgiu, Hofer and Hofmann 2014). For instance, buildings in an image could be classified across an image based on the brightness values of the objects representing them in the image. With GEOBIA, users obtain intelligence from the interpretation through a less subjective, less labor and time intensive process (Hay and Castilla 2008).

Though the military has likely begun utilizing GEOBIA for analysis, detailed documentation of processes remains classified for security purposes. Military applications of GEOBIA have developed noticeably in time. In the late 1990s, military analysts used simple, timely visual interpretation in the detection of the 1979 Indian nuclear missile test (Gupta and Pabian 1997). Recently, however, researchers adopted this "new paradigm" of GEOBIA to monitor specific sites of interest, namely for weapons treaty verification (Niemeyer and Nussbaum n.d.). These applications search for activity within facilities to identify suspicious activity suggesting weapons development. Given the recent Joint Comprehensive Plan of Action (JCPOA) in Iran regarding nuclear energy, GEOBIA research has focused on detection of suspicious activities in suspect nuclear sites. Another country which requires remote sensing-based monitoring of nuclear activity is the Democratic People's Republic of Korea (DPRK). Few unclassified studies aim to use GEOBIA in the DPRK to monitor its known missile testing facilities, however. A similar GEOBIA approach in treaty verification could be used in the DPRK to observe known missile site activities,

which would support the intelligence community (IC) in its monitoring for further developments in the DPRK missile program or for preparations towards another missile launch or test.

While the GEOBIA process inherently simulates human interpretation, it is unlikely that a single human can interpret every individual feature (image attributes such as spectral values, geometric parameters, etc.). A human interpreter is likely to focus on a handful of these features to distinguish a house from a tree, since he/she is cognitively incapable of utilizing the plethora of data available for a given group of pixels. To better mimic human intelligence of interpretation and improve classification accuracy, GEOBIA may exclusively limit its scope to those properties that the interpreter uses in his/her analysis.

This study aims to discover whether the use of interpreter knowledge for informing classification improves classification accuracy in the case of military feature identification. To accomplish this, here I compare the accuracies of classifying an image with exclusively humandetectable features to classifying with all available features. It additionally views images of the Palisades Nuclear Energy Facility in Michigan, US, as a spatial resolution comparison. This study uses the case of the DPRK's missile testing facilities for classification. The DPRK region is chosen in support of national security initiatives and to advance research towards monitoring of weapons of mass destruction (WMD). This study was conducted with the objectives in Table 1.

Objective 1 intends to gain an understanding of the human interpretation process and visual cues used for the analysis of images of missile test sites. To accomplish this objective, I needed to discover interpreters who write about their interpretations, particularly those who focus on the DPRK and its missile testing. I then extracted the interpretation elements used for the detection of each feature in the text. Further details regarding the process are provided in Section 3. I predicted that the contextual information from interpreters would focus largely on shape and texture elements and features, such as rectangularity and coarseness, translatable into a GEOBIA format. This hypothesis is based on the notion that much of visual interpretation of images involves the recognition of these interpretation elements and would likely accordingly apply to adding context to image analysis.

Objective 1	Obtain human interpreter knowledge.
Objective 2	Determine the features indicative of missile
	testing.
Objective 3	Segment and classify each of the images using
	the three classification methods and compare
	the accuracies with and without utilizing
	interpreter-used features.
Objective 4	Compare classification accuracies of multiple
	classification software.
Objective 5	Determine whether improving spatial
	resolution improves classification accuracies
	of the three classification algorithms.

Table 1: Objectives of the Study.

Objective 2 aims to pinpoint the objects of focus for detection and for assessment of the methods in Objective 3. Using the same interpretation rules from the content analysis conducted in Objective 1, I extracted the artificial features in the images that the interpreters highlighted as indicative of a missile testing facility or one which supports missile testing. I predicted that the

features that interpreters would focus on largely highlight different types of single-purpose buildings.

Objective 3 compares two cases. In the first, only image features that were identified in objective one were used in classification. In the second case, a full set of image features was used. I developed segmentations and classifications for each of the sites used in the analysis. For comparison, this step uses three methods of classification to determine if any combination of the classification and feature extraction are more accurate than the others. The methods of classification include rule-based classification, nearest neighbor, and random forest classification. I implemented all three classifications with both the complete feature selection and the interpreter-derived feature extraction achieved from Objective 1. An accuracy assessment was completed for each of the six classifications for comparison in detecting the objects discovered in Objective 2. I predicted that the use of the human interpreter elements would improve accuracies in detection across all three of the classification methods with the greatest increase in accuracy being in the random forest classification. I predicted that the random forest classifier would perform the best due to previous studies on the topic, discussed in the literature review of this thesis, as well as its inherent approach to the classification: the use of training and decision trees may better utilize different features for classification than the other methods based on the decision trees.

Objective 4 compares two different types of image classification software. I conduct in each software a random forest classification on the same study site using a knowledge-based reduced feature set between both platforms. The first software used is eCognition, which remains the primary software of this study, and the second is R (Trimble 2019; The R Foundation 2019).

Accuracy assessments were completed for each software and are compared for the classification of buildings in the respective image from Objective 2. I predicted that there would not be a significant difference in accuracies between the two software types with this method of image classification.

Objective 5 conducts an analysis similar to that in Objective 3 but instead compares the classification accuracies of two different spatial resolutions. The comparison will conduct all three classification algorithms used in Objective 3 for each of the two spatial resolutions of the same image. I did not use the same study site for this comparison as with the previous steps, but I did retain the parameters from the previous steps to isolate strictly the spatial resolution comparison. An accuracy assessment was conducted for each of the classifications of each spatial resolution, from which I could determine whether or not finer resolutions using these algorithms yield more accurate classification results. The accuracy assessments from this objective viewed the classification of all classes in the scene versus solely the building class as with the previous objectives.

The remainder of this document provides details, literature background, and results of the study. The following section details current and past research in GEOBIA, the incorporation of human cognition into GEOBIA, and the use of GEOBIA in cases of the DPRK. Section 3 discusses the data used in the study and the processes used to conduct the project in GEOBIA. Section 4 discusses the accuracy assessments produced in the project and all results obtained from the study. Section 5 discusses the results in the context of military intelligence and its potential implications and the limitations and delimitations of the project as a whole. Lastly, Section 6 concludes the project discussion and describes its contribution to future research and how a

similar project may improve the results of this project in future research. All acronyms and abbreviations used throughout this study may be viewed in Appendix 1.

## 2. Background

This study draws from history, geography, and Geographic Information Science (GISci). To understand the basis of the research, it is necessary to explain the context in which similar research has developed. Since the research maintains a military focus, I will first discuss military remote sensing, particularly for strategic operations. The studies and instances of remote sensing in this context involve both military-led and civilian-led research. Given this study's focus in the DPRK, a history of the nation's high-profile missile program, present capabilities and the research completed regarding the monitoring of this specific missile program are described. These details are necessary for understanding the pressing issue of DPRK to US politics and in gaining a relative perspective on the nation's motivations and testing capabilities.

Next, I review research concerning the application and development of GEOBIA. This explanation will exhibit the common methods of analysis in prominent GEOBIA studies, which contributed to decisions made in Section 3 of this study. I focus on rule-based classification methods to explain methods and applications common for this type of classification. I also review knowledge incorporation for GEOBIA. To provide further reasoning for our methods, we discuss trends in feature extraction and feature space reduction, objectivity in knowledge-based classifications, and potential workflow reusability.

Following a discussion of the past and present research in these relevant areas, I discuss the present research gaps and how this particular study may contribute towards these gaps.

## 2.1. Remote Sensing for Strategic Operations

The US and foreign militaries use GIS and remote sensing for a variety of applications in support of their operations. According to Witmer (2015), military operations were among the

first applications of remote sensing in violent conflict settings. Military remote sensing in the US began with the use of the hot air balloon to observe the battlefield in the American Civil War (Witmer 2015). These technologies allow observation of the battlefield from above and provide spatial information for military commanders, permitting these leaders and decision-makers to effectively plan from an aerial point of view (Satyanarayana and Yogendran 2013). Remote sensing technology has become increasingly popular with the US military, largely in the form of unmanned aerial vehicles (UAVs), since they do not put any lives directly at risk in the sake of reconnaissance.

Remote sensing research in support of strategic military operations takes the form of either digital mapping or UAV reconnaissance. Glade (2000) evaluates the use of UAVs in a variety of military applications, including transportation, intelligence and surveillance, attack missions, and combat support missions. For surveillance, the military has used the UAV technology for remote sensing reconnaissance since it remains relatively difficult to detect by those being observed. UAVs have also been used to remotely detect chemical and biological weapons autonomously (Glade 2000). These platforms have additionally been preferred by the military due to their ability to broadcast live information for long periods. Their use is particularly useful as it reduces the need to expose military personnel to fatigue and stress that operated flights cause.

Remote sensing in recent US conflicts has focused largely on the observation of urban environments, due to the military presence in both Iraq and Afghanistan. While urban environments are a major focus for remote sensing for current military operations, their difference from other environments and large variability (cities and towns vary in construction

techniques around the world) require smaller and less durable remote sensing platforms for operations (Samad, Bay, and Godbole 2007). The use of UAVs for surveillance emphasizes the military's situation awareness (SA) for understanding complex environments, such as the urban battlefield of today's conflicts (Samad, Bay, and Godbole 2007). Although UAVs and remote sensing are used for urban reconnaissance, military leaders also use the surveillance capabilities for detection outside the urban environment. Commanders use this view of the terrain to maneuver troops, materials, or vehicles, and to develop maps of the terrain for optimal resource utilization and decision-making for missions (Satyanarayana and Yogendran 2013). Due to the variety of environments, no broad ruleset exists for either military or civilian use in support of military operations or civilian research (Witmer 2015).

Military operations also use space-borne remote sensing platforms, such as satellites. The search for Osama bin-Laden, the mastermind behind the New York World Trade Center attacks in 2001, prompted the combination of Landsat 5 Thematic Mapper imagery and cultural geography to search for terrorist groups in the Zhawar Kili region of Afghanistan. This study resulted in the detection of terrorist posts containing terrorist-led convoys and potentially high-value targets in al-Qaeda (Beck 2003). Though military leaders use them in support of planning operations, space-borne platforms are not ideal for real-time military operations, due to revisit times over the same regions and the need for very high resolution (VHR) imagery, which is tougher to attain by space-borne platforms than from a UAV sensor (Witmer 2015). Maathuis (2003) accordingly declined the use of satellite imagery to detect individual landmines – due to the need for a much higher spatial resolution (in centimeters) than that achievable by airborne or spaceborne imagery. Due to the inaccessibility of this high-resolution data, this study instead

used Landsat imagery to detect the entire scene for the likely presence of minefields over a region (Maathuis 2003).

Military remote sensing also has a variety of targets. Witmer (2015) addresses the various methods and types of remote sensing used for detection and analyzing the effects of violent conflicts, including wars and genocide. Beck (2003) uses 30m Landsat imagery in combination with GIS and cultural geography to aid intelligence analysts in searching for terrorists hidden in mountains and caves in Afghanistan, while Maathuis (2003) used SPOT XS and Landsat Thematic Mapper imagery to detect minefields in three different regions of Zimbabwe to support civilian or military landmine clearance. All three of these studies use publicly available information in their military-oriented applications.

#### 2.2. The DPRK Missile Program

Some nations maintain the goal of secrecy, revealing as little as possible to the world outside their borders. Recently, the DPRK has emerged in the world news due to this secrecy and perplexing behavior on the global scale, particularly regarding its missile program. The DPRK poses a major threat to the US, making the development of an intelligent, making remote sensing developments concerning the missile testing sites of interest. The DPRK's ambitions for strategic weaponry rose almost immediately after its inception in 1950. Chinese and Soviet powers assisted the DPRK in achieving these aspirations. Between 1968 and 1969, the United Soviet Socialist Republic (USSR) provided a sample of its S-2 Sopka missiles to The DPRK for coastal defense (Sachdov 2000). At the same time, China provided similar assistance with its HY-1 naval missiles, themselves a bi-product of the USSR's SS-N-2 Styx missiles (Sachdov 2000). Additionally, Egypt provided the DPRK Scud B missiles in the late 1970s or early 1980s. Shortly after the

beginning of this weapons trading relationship, in 1972, the DPRK developed a domestic site to develop the Chinese HY-1 naval missiles.



Figure 1: DPRK Missile Ranges. Source: <u>https://www.dw.com/en/which-us-cities-could-north-koreas-ballistic-</u>missile-hit/a-39881831.

By the mid-1970s, the DPRK had made adequate progress toward its weapons ambitions with the assistance of its trade partners. In 1973, the DPRK possessed 24 unguided FROG 5/7 rockets and 6 SS-C-2b missiles (Sachdov 2000). The DPRK with China proposed the joint development of a single-stage tactical missile, DF 61, around this time. The project was canceled in 1978 due to the collapse of the main Chinese governmental supporters of the project, though the DPRK continued additional missile development. By 1981, the Korean weapons program accelerated with the cooperation of Egypt in a technological exchange agreement, which provided the DPRK Scud-B technology. The next year a North-Korean-built, Iranian-financed Scud-B missile was tested with three more in 1984. The DPRK established an official development and testing facility near the capital, Pyongyang, during the mid-1980s where it maintained an annual production of 50 Scud-B missiles (Sachdov 2000). The missile program and relative success for the DPRK have tightened relationships with Iran, in that the latter purchased in 1987 approximately 90-100 Scud-B missiles from the DPRK, due to its limited domestic assembly at its Isfahan plant (Sachdov 2000).

The DPRK missile program advanced in the late 1980s and early 1990s, climaxing in the development of the Scud C missile in 1987. The first test of Scud C occurred in 1990 and followed with full-scale production of 4-8 missiles per month over the next year. The program reached an elevated level when it developed the No Dong I missile in 1991, which was allegedly capable of reaching all of South Korea. The DPRK revealed this missile to other nations unfriendly to the U.S. – for instance, it attempted to sell it to Libya for \$7 million and exhibited it for Pakistani officials in 1992. DPRK weapons developers completed and finalized the No Dong I missile in 1993, testing two of them during missile testing from May 29-30, 1993 (Sachdov 2000).

The success of the Scud missile program in the DPRK inspired their nuclear weapons development. Though US sanctions have stunted the development of the nuclear program, the DPRK has continued to establish nuclear facilities across the country, most of which are located in Yongbyon. Currently, the Yongbyon facility possesses a 50 MW reactor, with other reactors across the country for testing and development. Sanctions from the U.S. froze the nuclear operations in the DPRK after the DPRK's attempted withdrawal from the Nuclear Non-Proliferation Treaty (NPT) in 1993 (Sachdov 2000).

The heightened focus on nuclear developments has accelerated the need for The DPRK to develop methods of delivery in the form of missiles. The first nuclear detonation test occurred at the Punggye-Ri underground testing facility on October 9, 2006, receiving plutonium from the Yongbyon nuclear facility (Chung 2016). Subsequent nuclear tests occurred at Punggye-Ri in 2009,

2013, and 2016, for a total of four tests. These tests have culminated at what the DPRK officials claim is the possession of a hydrogen bomb in 2016 (Chung 2016). As late as 2016, the DPRK has progressed towards the development of its KN-11 missile, which is a submarine-launched ballistic missile (SLBM) that can carry a nuclear warhead (Postol and Schiller 2016). The missile program has made significant increases in its capabilities since its withdrawal from the NPT in 1993.

Missile development by the DPRK has led to an increase in threatening political posturing by the DPRK towards the US. In 2016, in response to condemnations of potential hydrogen bomb testing, the Korean Central News Agency (KCNA) announced that the Iraqi and Libyan regimes of Hussein and Gaddafi, respectively, succumbed to destruction upon giving up their nuclear ambitions amid pressure from the U.S. and the Western nations (Chung 2016). The nuclear program came largely from the missile program, so it is worth focusing on the missile testing to combat the nuclear ambitions.

## 2.3. The DPRK Missile Development Monitoring

Although the DPRK is a national security interest for the US, not much public research has addressed methods of observing the DPRK's missile development with remote sensing. According to Shim (2014), the geographic study of remote sensing for The DPRK lacks focus and strength, regardless of the country 's abnormality as a *"terra incognita sui generis* or uncharted land of its own." Satellite imagery has been the focus for obtaining information regarding the developments in the DPRK due to lack of ground accessibility. The satellite imagery surveillance has yielded mixed results, due to classification inaccuracies resulting from image tampering or disguising objects on the ground to reflect a different object from space. Accordingly, greater details are obtained through independent monitoring services that may access the country on foot (Squassoni 2005). In its relatively small public domain, however, monitoring the DPRK for research largely revolves around non-remote sensing methods, or those not concerning its political aims and missile development. Since the DPRK's withdrawal from the NPT and denial of International Atomic Energy Agency (IAEA) inspectors from observing its plutonium enrichment facilities, remote sensing methods remain one of the only approaches to monitor the missile and nuclear development of the country (Pollack 2003). According to Albright and Brannan (2007), the IAEA provided independently *in situ* monitoring for the US intelligence estimates. Moreover, the monitoring focuses on the Yongbyon radiochemical facility, due to its plutonium enrichment and chemical laboratories on site.

Additional sensing monitoring of the DPRK activities from these sites takes the form of seismic wave analysis. Kim and Richards (2007) calculate the distance of seismic waves arriving at local monitoring sites and compare them to the times and locations of recorded earthquakes and compare these wave values with those from the nuclear test to detect the location of the test itself as well as the likely origin of the missile. Similarly, Schlittenhardt, Canty, and Grunberg (2010) estimated the seismic activity of the test site from different monitoring facilities and agencies for the DPRK test site, identifies the locations given these estimates in the low-seismic-yield region and confirmed testing activity using ASTER satellite imagery in a change detection analysis. This latter method differs from others that use seismic data for detection, as it chooses to complement the study and verify the missile tests with remotely sensed imagery. The other studies instead tend to focus on the seismic element created from the missile's impact and detonation. These studies address the monitoring of the DPRK for the Comprehensive Nuclear-Test-Ban Treaty (CTBT) without remote sensing data.

Several studies use remote sensing in the country to monitor activities at missile test sites. Albright and Brannan (2007) use of commercial imagery for monitoring and estimating developments of plutonium stocks at the Yongbyon radiochemical facility, and for monitoring related facilities at the same site. Broad et al. (2005) use satellite images by intelligence agencies and national monitors to detect tunneling activity likely used for missile and nuclear testing. Squassoni (2005) discusses the focus of satellite imagery for monitoring the DPRK by observing specific locations, namely the Yongbyon facility, and the components at these sites, such as the 5MW reactor at Yongbyon.

Some studies integrate remote sensing data with other information resources. Shim (2014) visually interprets nighttime NASA imagery to determine the spatial scarcity and development void in the DPRK in support of US policymakers and national security decisions. Ozeki and Heki (2010) used GIS to calculate a DPRK missile test's entrance and exit from the ionosphere using the channel frequency disturbances at various Japanese GPS station locations. This latter study focuses on the geography of Japan in conjunction with the geography of the missile launched from Musudan-ri, DPRK (Ozeki and Heki 2010).

Though these studies focus on identifying different components of the nuclear developments in the DPRK, none extracts the individual features of a missile testing facility. Furthermore, most studies use the Yongbyon facility due to its prominence in nuclear development and chemical enrichment. Minimal research appears to use the missile testing facility at the Sohae Satellite Launching Station, another prominent testing site.

#### 2.4. Rule-Based Image Object Detection

According to Castilla and Hay (2008), rule sets are a comprehensive representation of procedural knowledge. GEOBIA rule sets are used to classify the image based on user knowledge (Castilla and Hay 2008). GEOBIA is carried out through a multi-phase workflow: pre-processing the image, segmenting the image into candidate image objects based on spectral features, and classifying the image objects based on user-defined parameters for each class (Castilla and Hay 2008). Few studies in the reviewed literature stray from this workflow; although, some conduct a cyclical repetition of multiple iterative segmentations and classifications (Baatz, Hoffmann and Willhauck 2008). Castilla and Hay (2008) cite in Benz et al. (2004) that these cycles of segmentation and classification are necessary for the incorporation of semantic meaning into image objects.

Rule-based methods of image object detection in GEOBIA largely focus on the classification procedure, focusing mainly on either land cover classification or urban classification. Dragut and Blaschke (2006) classified the geomorphology of landforms in Germany and Romania by comparing Digital Terrain Models (DTM) to the image segmentation and classification based on a hierarchical classification rule-set. Bhaskaran, Paramananda, and Ramnarayan (2010) classified boroughs of New York, the US to compare pixel-based- and object-based classification methods. They address the separability of the urban features and focus on individual objects in the scene to create classes (Bhaskaran, Paramananda and Ramnarayan 2010). This focus on individual image objects serves as inspiration for the study described here. The authors conclude that GEOBIA significantly increased accuracies in detecting these urban features when compared to pixel-based methods (Bhaskaran, Paramananda and Ramnarayan 2010).

As a major component of rule-based GEOBIA, the feature thresholds and scales in the rules remain critical to the accurate analysis of the image using GEOBIA. GEOBIA allows for multiscale analyses at the pixel, object, and pattern levels. The object scale is the smallest meaningful unit for the analysis and most closely resembles semantically meaningful objects (Ming et al. 2015). Torres-Sánchez, López-Granados, and Peña (2015) investigate different values for the scale parameter, shape, and compactness in segmentation for classifying vegetation, determining that increases in scale parameter reduce error until an optimal value is achieved after which, the error increases. Through their implementation of multiresolution segmentation, the authors determine the scale parameter is the most significant parameter for initial segmentation parameters, since other values, such as shape and compactness, produced minimal impact in comparison (Baatz and Schäpe 2010; Torres-Sánchez, López-Granados and Peña 2015). Multiresolution image segmentation is a bottom-up pairwise merging technique that merges individual pixels with nearby pixels to produce an image object iteratively based on spectral similarity (Torres-Sánchez, López-Granados and Peña 2015).

Supervised methods include both k-nearest neighbor and random forest classification approaches. Supervised classifications in image analysis are those which assign classes to pixels or objects based on the spectral values from samples of each class. A common method for GEOBIA is a k-nearest neighbor, which Maxwell et al. (2015) used to classify mine presence and reclamation land in West Virginia, US compared to other methods of classification, such as random forests. In a K-nearest neighbor classification, the algorithm uses the samples to classify neighboring objects based on their similarity to the samples (Weinberger, Blitzer, and Saul, 2006). A commonly used machine learning algorithm is the random forests classifier. In machine learning, the processor collects spectral data based on a training set (often samples) and depending on the classifying algorithm, will use the information to train the classifier and classify the image. This method creates a designated number of decision trees and randomly plots the points in the image; classification is based on the training data at each location (Breiman 2001). Random forest classification has been used in various applications for land cover mapping, for example, peatland in Ontario, Canada using LiDAR data (Millard and Richardson 2015), urban classifications of LiDAR data (Chen et al. 2014), and determining tree health with IKONOS imagery (Wang, et al. 2015).

Rule-based classification studies tend to focus on incorporating knowledge into the image for interpretation. Krtalic (2016) uses a system, T-AI DDS, to conduct segmentation and classification for automatic detection of mines and minefields in Croatia. While the authors conclude that the application of automatically detecting mines requires more research, their semi-automated method of detection incorporates all data and expert knowledge available in the scene (Krtalic 2016). These types of methods follow a long history of expert-based image analysis systems. For example, Wharton (1982) proposed a knowledge integration method based on rules with the CONAN, or contextual analysis, method, which classified data into image components and conducted a contextual classification based on the mixture of these components.

#### 2.5. Knowledge Incorporation into GEOBIA Applications

With a focus on improving classification automation, studies in GEOBIA classifications are exploring the potential of complementing computational approaches with knowledge incorporation. Johnson and Xie (2013) developed a method to incorporate "super object information" into image objects and compare them with and without the information to

determine whether or not increased information improved the accuracy of detection. The information incorporated included the spectral parameters, texture, size and shape information of the image segments' super objects, which are the larger segments within which the segments of focus are located (Johnson and Xie 2013). Instead of comparing methods for accuracy, MacFaden and O'Niel-Dunne (2015) created two rule sets (one for each study site) and incorporated contextual criteria to detect automatically any potential infrastructure damage on roads following major storms using GEOBIA. Nussbaum, Niemeyer, and Canty (2006) similarly avoid comparison of multiple methods and develop a GEOBIA-based separability and threshold (SEaTH) algorithm for automatically classifying image objects. The study classifies objects in the Esfahan Nuclear Facility in Iran using a quantitative classification approach in the form of developed mathematic algorithms to classify the images (Nussbaum, Niemeyer and Canty 2006). This differs from previous studies that classify using qualitative features, namely the key interpretation elements (MacFaden and O'Neil-Dunne 2015). Marpu et al. (2008) use the same separability and threshold method and study site as Nussbaum, Niemeyer, and Canty (2006); however, their study uses this method to automatically process imagery in GEOBIA (Marpu et al. 2008). Unlike other studies, this study only focused on detecting a single class (the class of interest), which explained the investigation of the two classes, class of interest and background (Marpu et al. 2008).

The incorporation of knowledge into automated image detection processes has produced a large number of studies despite its complications. Many studies focus on knowledge incorporation via ontologies. Written, computer-readable, and reproducible representations of expert knowledge, ontologies are used in GEOBIA as a means of exploiting structural parameters of an image that is traditionally accomplished only by humans for image interpretation (Blaschke et al. 2014). Ontological workflow development is similar to standard rule-based procedures; however, developing workflows with ontologies involves hierarchies of classes for use and establishing a database of knowledge for defining individuals and classes (Gu et al. 2015).

Many studies centering on knowledge integration via ontologies use GEOBIA. Gu et al. (2015) studied the development of a universal workflow based on a hybrid machine-learning and semantic model, which they applied to farmland in China. Rather than using traditional GEOBIA software such as *eCognition* these authors employ web service applications (*GeoBrain*) to intelligently identify complex artificial features (Trimble 2019; Yue et al. 2013). This study focuses on the detection of weapons of mass destruction (WMD) facilities (Krtalic 2016; Nussbaum, Niemeyer and Canty 2006; Marpu et al. 2008). Belgiu, Hofer, and Hofmann (2014) developed a classification procedure with embedded expert knowledge using GEOBIA. After creating the ontology in *Protégé*, an open-source ontology building program, in the OWL2 Web Ontology Language, the ontology may be used to classify image objects (Stanford Center for Biomedical Informatics Research 2016; WC3 Web Ontology Working Group 2004; Belgiu, Hofer and Hofmann 2014).

Some studies have also examined the possibility of integrating knowledge into other image classification methods. For example, Liedtke et al. (1997) proposed a new program called AIDA that integrated semantic nets in image interpretation processes; however, the program could not successfully detect complex features as they were and could only detect these features. Similar to O'Neil-Dunne, MacFaden, and Pelletier (2011), this study attempts to replicate human interpretation in the creation and display of contextual relations, though this study focuses on

analyst queries whereas the former study addresses simultaneous segmentation and classification.

#### 2.6. Feature Extraction

GEOBIA is noted for having large computational demands. One means for reducing processing time and computer RAM demands is the use of feature space reduction, the process of reducing the number of image attributes used in classification based on target separability. A typical image object may retain more than 200 features across different scales. The number of object features increases as the spatial resolution becomes finer (Ma, et al. 2015). Limiting the number of features used for classification significantly reduces the amount of time needed to compute the classification, thereby making it effective for image analysts in need of quick results; however, it retains potential to decrease accuracy in classification given its fewer features of consideration (Ma, et al. 2015).

Land cover remains a common application for feature extraction methods in the present literature. Yu et al. (2006) classified local vegetation land cover in northern California, US using 52 features derived from a developed statistical classification and regression tree (CART) algorithm. A large number of object features is based on the CART algorithm, an automated approach with a higher number of features in this case than other studies (Yu et al. 2006). On the other hand, Taubenböck et al. (2010) focus on transferability from one classification to another using a limited number of object features. The study used an object feature hierarchy to classify and extract urban objects in Istanbul and India, which effectively classified individual homes at 85% overall accuracy (Taubenböck et al. 2010). Feature reduction affects several varieties of classification algorithms, including rule-based and machine learning approaches. According to Pal and Mather (2005), the size of a machine learning training set has a significant effect on the algorithm and must contain specific class descriptions. Machine learning feature space reduction must aim to include at least 10-30 times the number of features in each training set as pixels for training data (Pal and Mather 2005). Ma et al. (2015) analyzed the effects of training set sizes at different scales for rural land cover classification in Deyang, China. Using the gain ratio from Quinlan (1996), they rank all features and a best-first search algorithm to obtain object feature subsets (Ma et al. 2015). These reduced feature subsets and varying training sizes were used to conduct a random forest classification with minimal computational requirements (Ma et al. 2015).

## 2.7. Objectivity

A central concern in the integration of human knowledge into remote sensing workflows is objectivity. Objectivity in the detection and classification of image objects is a difficult task as human interpreters, as well as complex objects, are inherently subjective. Krtalic (2016) integrate computer-based segmentations with the self-produced T-AI DDS program to reduce the potential for subjectivity in the results. Furthermore, Gu et al. (2015) attempt to develop an objective GEOBIA workflow by incorporating expert domain knowledge via ontologies and semantic maps. Although GEOBIA classification is biased by an operator, subjectivity can be improved by using ontologies developed by multiple experts (Gu et al. 2015; Belgiu, Hofer, and Hofmann 2014).

Baatz, Hoffmann, and Willhauck (2008) attempts to disregard subjectivity concerns in developing a rule-based GEOBIA. Their approach requires the operator to know the real-world object for which he or she is searching in the image to segment the image and classify it, which retains some inherent subjectivity. The "spiral method" discussed and applied in the study requires the operator to segment by a particular item in the image, permitting the operator to choose when the segmentation and classification end and altering the scope based on different operator experiences in the domain (Baatz, Hoffmann, and Willhauck 2008). While the method retains inherent subjectivity in the knowledge of the features for segmentation, it attempts to establish a definition based on the parameters that are useable by different operators for classification. Moreover, Arvor et al. (2013) argue that in GEOBIA applications, expert knowledge and expert bias ultimately limit the segmentations and classifications, emphasizing the need to reduce subjectivity and increase objectivity in GEOBIA workflows.

#### 2.8. Workflow Reusability

A universal rule-set or workflow is unlikely to be achieved using GEOBIA due to the complexity and heterogeneity of the earth's surface. Creating limited, repeatable workflows for specific scenarios may be possible. Ontologies are one means of addressing this issue. At present, there is a lack of a comprehensive, systematic formalization for class definitions in GEOBIA leading to subjectivity (Belgiu, Hofer, and Hofmann 2014). These authors continue to address that ontologies may potentially serve as standardization for class definitions (Belgiu, Hofer, and Hofmann 2014). Arvor et al. (2013) explore the potential for ontologies and argue that they will permit objectivity across disciplines, emphasizing ontology mapping as a means to that end. This sentiment is shared across other studies (Castilla and Hay 2008).

Despite the potential for a universal approach and workflow for GEOBIA applications, few follow through and conduct testing towards neutral and reusable workflows. Yue et al. (2013) incorporated thematic semantics to develop a transferable workflow for weapons site detection.

Studying farmland in China, Gu et al. (2015) pursued not only a fair, objective workflow but also one that would be universally applicable based on machine learning methods. None of these studies has led to the development of a successfully transferable GEOBIA workflow (Gu et al. 2015).

#### 2.9. Contributions of the Present Research

I have reviewed some research studies that have addressed various facets of GEOBIA workflow. While these studies have compared outcomes from the application of different classification strategies, few of these studies address the usefulness of incorporating expert knowledge. A variety of knowledge incorporation methods, such as ontologies and semantic networks, have been used in GEOBIA. Feature space reduction appears to be a more common approach to classify image objects intelligently. As far as I can tell, no study directly compares the use of feature space reduction as a knowledge-based process to a classification without feature space reduction. Other studies also do not appear to have used expert interpreters' annotations as a guide for feature selection. Thus, our research explores a new avenue for feature space reduction and potential for improving the integration of expert knowledge into GEOBIA.

Despite a limited number of studies on knowledge integration that focused on military sites, none focused on the classification of DPRK missile testing facilities, which is surprising given its prominence in modern global affairs. Only a handful of studies focus the remote sensing imagery on missile testing sites. GEOBIA research appears to favor of the Islamic Republic of Iran (IRI) and India (Niemeyer, Marpu, and Nussbaum 2008; Gupta and Pabian 1997). Though these studies focus on the CTB treaty verification, they do not address the issue in the DPRK's missile program, at least publicly. Finally, there does not seem to be a systematic approach to automating GEOBIA
for commercial intelligence uses. According to Diamond (2001), the lack of quasi-instantaneous automated change detection in satellite imagery as a major problem for government intelligence. This study will fill the gaps left regarding the lack of research towards the DPRK missile program in GEOBIA and a direct comparison of automated methods of classification comparing knowledge-based- and knowledge-devoid-classification methods.

# 3. Methods

The goal of this research is to develop a GEOBIA workflow for identifying buildings at missile test sites in the DRPK and to determine whether the inclusion of expert knowledge improves the object detection accuracy. I conducted four main analyses to meet the five objectives of this research.

**Objective 1**: Obtain human interpreter knowledge.

**Objective 2:** Determine the features indicative of missile testing.

**Objective 3:** Segment and classify each of the images using the three classification methods and compare the accuracies with and without utilizing interpreter-used features.

**Objective 4:** Compare classification accuracies of multiple classification software.

**Objective 5:** Determine whether improving spatial resolution improves classification accuracies of the three classification algorithms.

The first analysis extracted expert interpreter's visual cues used for the detection of missile testing facilities from text documents. The second analysis compared the different GEOBIA classification methods to determine which classification techniques prove more effective in classifying buildings of the missile testing facilities. The third analysis then compared these classification techniques across classification software. Finally, the fourth analysis compares the classification algorithms across different spatial resolutions. Traditional remote sensing accuracy assessment methods are used throughout the study to determine the success of these classifications.

### 3.1. Study Sites

The study areas of this project include two major missile-testing facilities in the DPRK, as well as the Palisades Nuclear Energy Facility in Michigan, US. Data for the DPRK sites, their locations, and test information were retrieved from the Nuclear Threat Initiative (NTI) website (Nuclear Threat Inititative 2018). The NTI compiled a Microsoft Excel spreadsheet containing missile test information since 2016 with variables including date of tests, type of missile tested, the location of the launch, the achievable distance by the missile tested, and the test's success or failure. For this study, I used the test site location from this dataset.

The DPRK monitoring agency, 38 North, publishes web articles containing very highresolution satellite imagery from Digital Globe Company along with interpretations by former military analysts in the form of image annotations (38 North 2018). I collected all articles and associated images from 38 North's satellite imagery archive through February 2018, which totals roughly 120 articles with approximately 1500 images across all of the articles. Each image contains between one and ten analyst-written annotations to highlight specific components critical to their analysis of the testing sites, such as a launch tower arm or vehicles present or absent.

The articles were then grouped by the missile test site location, and the five sites with the most articles were chosen for the study. Though the DPRK's main nuclear testing site, Punggye-Ri, had the greatest number of articles, I excluded it as an outlier being the only confirmed nuclear weapons testing facility in the DPRK (Nuclear Threat Initiative 2018). Therefore, I selected the next four most-frequently-occurring sites from which to analyze and increase the focus of the study-- Sohae, Yongbyon, Sinpo, and Musudan-ri (Tonghae). I selected this subset for several

reasons. First, the four sites provided wide spatial coverage horizontally across the country, which would ideally require a more comprehensive rule-set that accounts for broader geographic variation in the land surface, cover, and uses. Second, the additional sites are seldom mentioned in 38 North, making them less appropriate for this study given their likely infrequent use. Lastly, the use of four study sites allows a manageable amount of data to analyze and compare with our methods as opposed to data from each of 23 testing facilities.

The four sites yielded 300 total images for which I created an Excel spreadsheet to compile each image and article's data, including the site focus of the article, date of the article, the name of the interpreter(s)(left as 38 North if none other indicated), the date of the image, and the annotations in each image. I assigned a specific identification (ID) code to each image and its respective data and qualitatively coded the annotations, which will be discussed in detail in Section 3.2. Figure 2 displays the number of articles and interpretations per the four missile testing sites. After further assessing the study sites, I chose to focus on only the Sohae and Yongbyon facilities. Imagery available for Sinpo and Musudan-ri did not permit adequately fine resolution to distinguish the individual building facilities. Additionally, these two omitted sites operate quite differently from Sohae and Yongbyon, which are both largely urban compared to the shipyard, Sinpo, and mountain launching facility, Musudan-ri.



Figure 2: Number of expert interpretations from 38 North per missile testing site in this study.

I justify the selection of the Sohae and Yongbyon sites because these sites still adequately address urban areas in the DPRK and are likely reproduceable at other similar study sites. Second, these two remaining sites retained far greater articles of focus than the respective omitted sites. Lastly, focusing on two sites permits far greater focus into the study applications than spreading resources across the four sites. For both Sohae and Yongbyon, I retrieved imagery for both locations from Planet Labs Imagery & Archive (Planet Labs 2019).

## 3.2. Data

The imagery used retains a 1m spatial resolution with four spectral bands and was acquired in 2016. For the analysis of the different spatial resolutions, the image used was obtained from the US Geological Survey (USGS) Earth Explorer as National Agricultural Imagery Program (NAIP) imagery. The image retains a 1m resolution initially, which was degraded to 3m resolution for the comparison. This image contains four spectral bands and is 1,936 by 2,526 pixels in size. All of the original images used may be seen in Figures 3-5.



Sohae Satellite Launching Facility, DPRK

Figure 3: Sohae Image. Source: Planet Labs

Yongbyon Radiochemical Facility



Figure 4: Yongbyon Image. Source: Planet Labs.

Michigan Nuclear Power Plant, 1m

Michigan Nuclear Power Plant, 3m



Figure 5: Palisades nuclear power plant, 1m resolution and 3m resolution (right). Source: USGS.

#### 3.3. Analysis 1 - Extraction of Expert Interpretation Cues

The objective of this study was to extract expert knowledge from image interpretations on the 38 North website and analyze the contests of those interpretations using content analysis. Content analysis is often used for qualitative research, a process where text contents are grouped into classes based on a theme relevant to the study (Hsieh and Shannon 2005). In content analysis, all textual information is classified based on a case-based dictionary of terms (McTavish and Pirro 1990).

I use content analysis to codify the expert annotations. After collecting all of the experts' annotations, I codified the annotations hierarchically based on common characteristics. I transformed the annotations manually in Microsoft Excel into single-word codes using descriptive coding, which provides a general descriptive term for a breadth of information (Saldaña 2013), and *a priori* coding, which employs theoretically-derived codes from prior knowledge (Bazeley and Jackson 2013). For example, buildings and train stations received the code "building." I then divided these general codes into more specific categories in a similar *a priori*, descriptive manner. In the same case as above, train stations received the class, "building," and the subclass, "transportation." I further extracted the annotations and organized them in an Excel spreadsheet. In the end we identified six classes: buildings, vehicles, environment, missiles, and changes. Since the interpreters of the images used repeated the annotations across images due to similar characteristics in different images, I omitted any repeats in the organization of these classes for individual annotations.

Classification of these text documents resulted in a data dictionary that could be used to support the second objective of this work. The "buildings" class yielded the highest number of

35

annotations and objects after eliminating duplicate mentions. The remaining classes were merged into one of three classes: vegetation (environment), water (environment), and built-up (all unaccounted annotations).

After determining which objects were of most interest, I performed a text analysis of the expert-written articles. This analysis aimed to determine the words, and thus visual cues, expert analysts used in interpreting the images. I employed the online corpus developer, *Sketch Engine*, to create a corpus and determine these contextual cues. *Sketch Engine* allows the user to provide a corpus of text and extract data about words, parts of speech, position in a sentence, and so on (Lexical Computing CZ 2019). I first created a corpus including all of the articles from 38 North about the Sohae and Yongbyon study sites for this analysis.

From these articles, I then analyzed the corpus using the image interpretation elements of texture, shape, size, pattern, shadow, location, tone/color, height, and site as search filters. Each of the searches was conducted individually and included related terms. For instance, the analysts did not often write the word, "texture" or "color" in their reports; however, they would write "smooth" and "green," which thereby yielded greater results than the interpretation terms. These results are similar to the results of the content analysis of historical interpretation documents presented by Bianchetti and MacEachren (2015).

After the word frequency analysis was completed on the article corpus, a concordance analysis was carried out on the interpretation elements in the document. The entire concordance results consisted of 305 entries, retaining the searched word, its grammatical form in the sentence, the sentence fifteen words before and after the cue, and the document it in which it exists. I then reduced the concordance to a subset which included only the searched adjective

36

cue, its frequency across the articles, and its cognitive code relating to the image interpretation elements. Table 2 displays the concordance subset for identifying buildings in images. These most-frequently-occurring textual cues were then used in the second portion of this study to incorporate expert knowledge as a means of feature space reduction in image analysis. For example, what would be the best features for identifying buildings? The "buildings" class yielded the highest number of annotations and compressed (after eliminating repeats) objects.

<b>Adjective</b>	Frequency	Cognitive Code		
green	5	Color		
high	32	Height		
low	26	Height		
short	16	Height		
rectangular	7	Shape		
large	107	Size		
small	96	Size		
long	14	Size		
Table 2: Concordance subset for the knowledge base				

 Table 2: Concordance subset for the knowledge base.

## 3.4. Analysis 2 - Classification Methods Most Appropriate for Missile Facility Extraction

Next, GEOBIA was carried out to determine which classification method performed best at identifying buildings. Trimble *eCognition* was used to develop the GEOBIA workflows, resulting in six image classifications per each of the two images. Trimble's *eCognition* simplified the translation of the concordance results in the image classification.

Three of the six classifications included the results of the concordance analysis to reduce the feature space. The other three classifications used the same type of classification method but did not include a feature space reduction method. The three classification methods selected for analysis were rule-based, nearest neighbor, and random forest classification. The rule-based classification requires the development of rules based on thresholds of spectral and geometric information to classify the image objects. These rules require the user to manually determine the appropriate thresholds for the features used to classify the image. As such, this method required the most trial-and-error iteration of the three methods used here and was the most timeconsuming.

The two supervised approaches, nearest neighbor and random forests, used in the study required training data to classify the images. The nearest neighbor classification method uses the feature values from samples and determines, based on the similarity between the training classes and candidate objects, which class is most appropriate for each candidate image object neighboring the samples. It evaluates the similarity of each object to its neighboring class and assigns a class based on whether or not it is more similar to that given class.

The random forest classification is a common method of machine learning classification. Similar to the nearest neighbor classification method, the random forest classification requires a training set. This layer of samples is used to train the algorithm to detect the assigned classes based on the value thresholds of the class samples. The algorithm creates a user-designated number of decision trees for the computer to process, and a user-defined number of random points. After the algorithm is trained, it creates many trees, based on the user's discretion, and determines the classes at each tree, after which it conducts a vote across all trees that determines the class at that location.

These classification methods were then augmented to incorporate expert knowledge into the classification. Feature space reduction decreases the number of features, or attributes, used for classification. This approach tends to improve computation speed and improve accuracy, as some features used for classification remain difficult or impossible for an expert interpreter to employ in analysis and may introduce noise into the classification process. Table 3 displays the

38

segmentations and classifications for the computer knowledge (CB) and expert knowledge (EK) feature spaces in the study.

<u>Computer Knowledge</u>	<u>Expert Knowledge</u>	
Rule-based	Rule-based	
Nearest Neighbor	Nearest Neighbor	
Random Forest	Random Forest	
Table 3 Classifications used for analysis per feature space.		

### 3.4.1. Rule-Based Classification in eCognition

To initiate the study and develop a relative understanding of the thresholds of values in the data, I began with a rule-based classification. The rules used in the classification are developed based on user-designated thresholds of the spectral values of the image objects. Manually inputting these values provided us a better understanding of the values in the data for the future classifications.

Before developing the rules for the classification, I first had to segment the image into candidate image objects. I used the same segmentation parameters for both images. I used a multiresolution segmentation parameter of a 50 scale, 0.1 shapes, and 0.5 compactness for both segmentations. These values were selected based on an interactive process and visual interpretation. Following the segmentation, we conducted a merge of image objects with similar spectral values. I manually merged the image objects by hand using *eCognition*'s manual merge tool to merge similar objects, focusing the image objects representing buildings.

After the segmentation and merging, I began to develop the classification rules. Each rule focused on classifying a particular class. I classified the most prominent classes first, progressing to classes with fewer features to classify. I used the following classes for both images: water, vegetation, built-up, and buildings. A background class was created to delineate the void portions

of the images due to the path of the satellite from the actual image content. Since all values in the background class were 0, I classified this class first, followed by the four classes in the order listed above.

To determine the thresholds, I focused on a select number of features to build the rules. I used *eCognition*'s feature selection window to test different thresholds of particular features through a trial-and-error method of inputting values. For the CK classifications, I did not limit the number of features for analysis. Accordingly, the rules accounted for all four spectral band values, geometric values, and shape and positioning values, in the CK case, while the EK classifications considered mainly spectral values and geometric values. While I considered these features in the development of rules, I developed rules mainly on the spectral values of the classes, including brightness, near-infrared, red, blue, green, and maximum difference values. For the EK rule-based classification, I limited the feature space used for rule development to only those found useful in the text analysis. The features in this second case were limited strictly to the most equivalent feature in *eCognition*'s feature space to the word in the concordance. A comparison of the features used in each method can be seen in Table 4.

	CK Features	<b>EK Features</b>
Brightness	The radius of Smallest Enclosing	Brightness
	Ellipse	
Max. Diff.	Rectangular Fit	Green
Blue	Roundness	NIR
Green	Shape Index	Area
NIR	Number of Pixels	Number of Pixels
Red	Border Index	Asymmetry
Area	Compactness	Elliptic Fit
Border Length	Density	Rectangular Fit
Rel. Border to Image Border	Elliptic Fit	Roundness
Volume	Main Direction	
Asymmetry	The radius of largest enclosed	

Table 4: Features that were used for knowledge incorporation.

The rule development process continued until I determined the classification produced the best results. In eCognition, I produced these rules in a hierarchy that separates the segmentation rules from the classification rules and separates the rules for each class. In doing so, the classification process can be reapplied without having to re-segment the image. The classification parameter values varied between the images, but I retained the rules between the same images. The final rulesets for each of the CK and EK classifications are displayed in Appendices 2a-2b. I repeated these classification iterations with different rule thresholds until the best-appearing classification was produced.

## 3.4.2. Nearest Neighbor Classification in eCognition

The nearest-neighbor classification was completed next. For consistency in the segmentation, the same segmentation parameters were used as those described above. The supervised nearest neighbor approach does not require specified rules; however, it requires a set of samples to compute the classes itself. Using the feature space designated in Table 4 above for the without-knowledge classifications, I created samples for each class by visually sampling objects for each image using the sampling brush in *eCognition*. While I attempted to keep the number of samples the same across both images, the significantly greater number of image objects in the Yongbyon image led to a slightly greater number of samples in this image than in the Sohae image.

Additionally, the number of samples varied across the classes, since some classes, were more easily distinguished. The number of samples per class in each image are seen in Table 5.

Class	Sohae	Yongbyon
Destidies a	25	25
Building	25	25
Background	1	1
Water	6	10
Vegetation	31	25
Build-up	20	30

Table 5: Number of samples per class for each DPRK image.

The number of samples used was scaled to the amount of the respective class objects in the image, with a greater emphasis on the building objects and samples. The samples were held consistent for both the case that included feature space reduction and the case that did not. The samples for both images are visible in Figures 6-7.



Figure 6: Sohae samples used for supervised and machine learning classifications.

For both the knowledge-based nearest neighbor classification and without-knowledge-based variant, I used the same segmentation parameters used in the knowledge-based rule-based



Figure 7: Yongbyon samples used for supervised and machine learning classifications...

classification and used the same samples for both nearest neighbor knowledge levels. The same image objects were accordingly analyzed for the algorithm tests across both feature spaces of the same respective images. The difference between the two tests per each image is the different feature spaces used to classify each in the nearest neighbor algorithm. The knowledgeincorporated nearest neighbor classification, I used the knowledge-based feature space to produce its supervised classifications. The without-knowledge variant, however, used the nonrestricted feature space representing no knowledge incorporation. The same two rules and the same classes were used for both with- and without-knowledge classifications of both the Sohae and the Yongbyon images to maintain consistency between the study sites. This algorithm produced a total of four classifications, as with the other algorithms in this study.

#### 3.4.3. Random Forest Classification in eCognition

The final classification algorithm is a random forest classification. The segmentation parameters were held constant for the without-knowledge and the knowledge-based classification workflows resulting in the same candidate image objects as the rule-based and nearest neighbor classifications above.

This classification required three rules: the image layer copy, training the classifier, and executing the classification. To train the classifier, I used samples transferred from the nearest neighbor classification. The features used for the training included those in Table 2 from the feature spaces. As before, the feature spaces used differed between the without-knowledge and the knowledge-based processes. Within this trainer, I also designated the parameters for the random forest classification. I used 16 maximum categories, 200 trees, and 0.2 forest accuracy to produce the best-appearing classification from the random forest classification. All other parameters in the editor remained their default settings.

Following the establishment of the training parameters, I trained and ran the machine learning classifier. Executing the classifier required the use of the final rule in the ruleset. All parameters for this rule were default parameters except selecting the appropriate training data. I conducted the same steps with the knowledge-based random forest classification with only the reduced feature space. As with the other classifications, the random forest classifications for both images were rerun until the best-fitting parameters were achieved.

#### 3.5. Analysis 3 – Comparison of Classification Software

Several programs exist to conduct image classifications. While this study primarily uses *eCognition*, to find the combination of variables that produces the best classifications of the

44

missile testing facilities in the DPRK, I produced classifications using the *R* software package. This package permits the use of the *R* computer language to conduct the classifications manually. To compare the software packages, I conducted a random forest classification in *R* on the Sohae image to compare to the results obtained from classifications conducted in the previous analysis.

Both *eCognition* and *R* have capabilities for executing random forest classifications, though their purposes and processes remain quite different (Trimble 2019; The R Foundation 2019). *eCognition* is a GEOBIA software developed by Trimble that permits the development of rulesets to automatically classify and analyze remotely-sensed imagery. *eCognition Developer* is used to develop processes for image analysis whereas the other packages extend the software for other analysis scenarios. *R* is a highly flexible, open-source, and extensive statistical program used for a variety of applications. It is a language that relies on coding to produce some statistical operations in various programming environments. The wide applicability of the *R* computer language permits the analysis of remote sensing data and conducting classifications of its own.

For the comparison of the software techniques, I used the random forest classification algorithm in *eCognition* and *R*. I used only the Sohae image for the comparison using the withoutknowledge case to retain consistency between the two programs. To conduct the classification in *eCognition*, I used the same steps detailed in Section 4.4.3. All of the parameters detailed in 4.4.3 discussion of the without-knowledge case apply here.

The process of conducting random forest in R does not require workflow development in the same way that eCognition does, but rather, it uses command line processing to call on packages to classify the image. To train the random forest classification, a reference dataset constructed in ArcMap resembling the samples used in the eCognition classification was used. Using the *R* 

*ranger* library, I conducted random forest classifications and constructed three separate models. I began by creating a point sample using the reference data and the pixel values from the image. Of the 69 polygons in the reference sample points, 50 were chosen at random for sampling. Within the chosen polygons, three-point locations were chosen at a regular sampling interval, removing points within 15 meters from another, resulting in 139 points.

Spatial random forest (*RFsp*) classifications were executed in *R*, by generating a grid of 29 points across the image, with one raster generated for each point. Cell values then represented the distance from the cells to each point, which enabled the classifier to use the relative location for improved predictions. Of the 139 sample points, 111 were used for the training data, and 28 were used as a validation set.

With these training and validation sets, I constructed three models with the *RFsp* classifier. The first model used the pixel value for the random forest classifier. The parameters designated were trees=500 and *mtry*=one, *mtry* is the parameter that constraining the number of variables considered at each decision point in the random forest classification. The second model developed and used in R is as follows with equation (1):

(1)

where z is the pixel value.

Each *layer.x* value represents the distance value from each respective grid point. The model identifies the best *mtry* parameter value. Once *mtry* was defined, the classification was run with

500 trees. The third model did not consider pixel value, as it was a purely spatial *RFsp* model, and the model used in R is as follows with equation (2):

```
Class = layer.1 + layer.2 + layer.3 + layer.4 + layer.5 + layer.6 + layer.7 + layer.8 + layer.9 + layer.10 + layer.11 +
layer.12 + layer.13 + layer.14 + layer.15 + layer.16 + layer.17 + layer.18 + layer.19 + layer.20 + layer.21 +
layer.22 + layer.23 + layer.24 + layer.25 + layer.26 + layer.27 + layer.28 + layer.29
```

(2)

The *mtry* value used in this third model was 10, and again, 500 trees were used. The results of each model will be displayed in Section 5.

## 3.6. Analysis 4- Comparison of Spatial Resolutions

After achieving poor results using the Planet Lab imagery, I examined whether to improve the spatial resolution of the satellite imagery might improve the classification results. Because imagery of the DPRK was not available at resolutions greater than 3 meters, I used proxy imagery for this final analysis. The three-meter resolution made a visual interpretation of the objects in the image difficult.

I chose to use NAIP (National Agricultural Imagery Program) imagery with a 1-meter spatial resolution of the Palisades Nuclear Energy facility in Michigan, US as a proxy for the Planet Labs imagery, as it has a finer spatial scale, but similar spectral resolution (United States Department of Agriculture 2015). A 2016 NAIP image of this site was obtained from the US Geological Survey (USGS) Earth Explorer (United States Geological Survey 2019). Insets of the image of the nuclear energy facility are provided in Figure 8 to display in more detail the features of focus. For consistency and a further test of the accuracy between different spatial resolutions, I reduced

the resolution of the image in ArcMap to three-meters to imitate the spatial resolution of the Planet Labs imagery of the DPRK missile testing facilities.

I conducted the same six classifications described above including the distinction between without-knowledge and knowledge-based feature spaces. I copied the exact rules used in classifying the Sohae image and employed them towards the classification of both the one- and



Figure 8: Nuclear Power Plant (top) and airfield (bottom). 1m images on left and 3m images on right.

three-meter resolution images of the Michigan nuclear energy facility. The segmentation used the same parameters as the Sohae segmentation. The parameters of the rule-based classification required altering given the difference in the spectral parameters of the Planet Labs imagery to the NAIP imagery. Rules for the other classification algorithms were not altered to maintain as much consistency as possible between the images.

The classifications of the NAIP images followed the same steps for classification of rule-based, nearest neighbor, and random forest classifications. The aim of this classification was to create a full-scene classification as opposed to focusing on the buildings present. Some classes were added – namely, the residential and industrial buildings classes to distinguish the two types of buildings visible in the image and determine if the resolution permitted such analysis. Due to this addition of classes, some new rules were required for the rule-based classification; however, several of the rules remained similar to those copied from the Sohae image classification. The complete rules are found in Appendix 2c.

The new image and new classes required new samples for the nearest neighbor and random forest classifications. The samples from the Sohae imagery would not be representative of the Michigan nuclear facility either in geography or in spectral values. I created samples for each of the new classes: residential buildings, industrial buildings, built-up, vegetation, and water. The background class was excluded from the classes since the image did not contain any spectrally void regions. The numbers of each of the samples are as follows: 25 residential buildings, 15 industrial buildings, 20 built-up, 30 vegetation, and 10 water. The samples were the same for the without-knowledge feature space, but the samples changed between the two resolution images. Due to the different resolutions, the segmentation algorithms produced different results with far fewer image objects in the three-meter resolutions than in the one-meter resolution. As a result, the samples from the one-meter resolution imagery could not be transferred accurately to the three-meter imagery, despite applying a Test Time Augmentation (TTA) mask for sample creation. Applying the TTA mask required a degree of overlap between the two images to be successfully transferred. In doing so, the default overlap of 75 % produced no samples in the new image. I reduced this number incrementally to 10%, which ultimately resulted in very few samples in the new image. Accordingly, I chose not to use the TTA mask for the sample selection of the three-meter resolution image and instead created samples resembling as closely as possible the samples from the one-meter imagery. The samples are displayed in Figure 9 to show the differences in the samples' sizes, though the distribution and number remain relatively similar.

The nearest neighbor and random forest classifications retained the same rules and parameters aside from different samples between the resolutions. I ran each algorithm until I achieved optimal results. The residential building and industrial building classes were removed from the classifications in the three-meter resolution imagery due to major misclassifications. This issue is discussed in depth Section 5. As with the other classifications, the results of each of these classifications is provided in Section 4. comparison. Each of these types of accuracy assessments reflected the nature of the information sought from each classification.

50



Figure 9: Samples used for the spatial resolution comparison classifications.

## 3.7. Post-classification Accuracy Assessments

The comparison of several classifications with several different factors required more than one type of accuracy assessment. Classification accuracy assessment was conducted for each of the classification cases that are described above (DPRK imagery, NAIP imagery).

Accuracy assessment of the DPRK imagery was completed first. The goal of this assessment was object identification accuracy. First, all of the Sohae and the Yongbyon images classification were exported as shapefiles to ArcMap. I created a shapefile for each image of the features identified by the 38 North image interpreters. I then created a random points layer for each image. The attribute table of this random points layer was then joined with each of the six classifications results shapefiles. These joins resulted in a file that contained the class results and the expected results based on the expert interpretations.

To assess the accuracy of the DPRK image classifications, I manually created a confusion matrix from the attribute tables. For each classification, I visually observed the classes of each

random point based on the respective unclassified images. This visual result acted as the ground truth for the accuracy assessment since I have limited access and resources to provide a more detailed ground truthing. I then compared the visual interpretations to each point's respective classification in each of the six classifications per image.

I labeled the classified points as either building or non-building. I included the building of interest points created to determine how each classification algorithm performed. The building points were compared to their classifications in each algorithm. For each algorithm, out of the total number of image objects, I determined the number of true positives (TP) or correctly classified buildings; false positives (FP), or non-buildings classified as buildings; true negatives (TN), or non-buildings classified as non-buildings; and false negatives (FN), or non-buildings classified as buildings. With these values for each of the classifications, I calculated the user's (UA) and producer's accuracies (PA), the overall accuracy, and the value of F (Radoux et al. 2011). The equation used to calculate each of these values is as follows:

CA = [(TP+TN)/(TP+TN+FP+FN)], UA = [TP/(TP+FN)], PA = [TP/(TP+FP)], F = [(UA x PA)/(UA+PA)

(3)

Next, I developed an accuracy assessment to compare the results of the *R* and *eCognition* processes. To compare the different software, I needed to compare the accuracies of the processes from both the eCognition random forest classification and that produced in R. I calculated the values for identification accuracy assessment as detailed above for both the

eCognition random forest classification (without knowledge) and the *R* random forest classification using the four classes used in the classification, eliminating the background classification altogether. The values determined for the R classification used the validation dataset created using the 29 generated points and compared the locations of these points to the classifications for each model created. Contingency tables were created for all models present between the two classifications.

The final accuracy assessment compared the impact of spatial resolution on the accuracy of NAIP classification. Since these classifications focused on classifying the entire scene of each of the images, rather than extracting building features, I constructed a confusion matrix of each of the classes. The steps to create the table follow a similar method for extracting resulting classes and reference classes. I created 100 random points and linked these points to each of the three classifications for each spatial resolution. This link assigned each random point the class present at its location in each of the classifications. A contingency table comparing the visuallyobserved location of the random point in the image to its respective classification was created for each classification for each of the two images. For validation consistency, I retained the same random points for each of the contingency tables, only changing the classifications.

53

# 4. Results

The goal of this research was to determine whether the introduction of human knowledge into the GEOBIA process could improve image classification accuracy in the case of missile test sites. Due to a lack of success in the classification of DPRK images, a secondary analysis was conducted using USGS NAIP imagery as well. This section provides the details of classification results of the three comparisons in this study. Processing times for all classifications may be seen at the end of Sections 4.1 and 4.3 in Table 7 and Table 18, respectively.

## 4.1. Knowledge Incorporation Comparison

This section provides the results of the knowledge comparison classifications. I provide results of each of the rule-based, nearest neighbor, and random forest classifications for both the withand without-knowledge feature spaces. Each subsection provides details for both the Sohae and the Yongbyon images and is detailed accordingly. For each of these classifications, I used the TP, FP, TN, and FN to calculate image object detection accuracy (overall, user's and producer's accuracy), as discussed in the previous section.

# 4.1.1. Rule-Based Classification without Knowledge in eCognition

To assess the amount of time that it took to classify the images, I accounted for the number of rules needed in lieu of time. For the Sohae image, the ruleset required a total of 10 rules. The resulting classification for the Sohae rule-based classification without knowledge is seen below

Rule-Based, No Knowledge



Figure 10: Rule-based no knowledge classification, Sohae.

Mile

# Rule-Based, No Knowledge



Figure 11: Rule-based no knowledge classification, Yongbyon.

Of the 31 image objects representing buildings in the unclassified image, 25 were classified correctly as buildings with 6 were misclassified. A table of the accuracies and true/false positives/negatives may be seen with according accuracy calculations in Table 6 at the end of Section 4.1.

For the Yongbyon study site, I used a total of 8 rules. The representative classification of the Yongbyon site with this classification may be seen below. Of the 27 building image objects, 13 were classified as buildings and 14 were misclassified as other classes. One image object was incorrectly classified as a building (FP).

# 4.1.2. Rule-Based Classification with Knowledge in eCognition

The second rule-based classification followed the same methods as the previous rule-based classification, only with a reduced feature space based on knowledge from image interpreters. For the Sohae image, I used 6 rules. Of the 31 image objects, 25 were correctly classified as buildings and 6 misclassified. For the Yongbyon image, 10 of the 27 were correctly classified as buildings, with one false positive. I used 5 rules for this image classification. The respective classification maps are as follows:



Rule-Based, With Knowledge

Figure 12: Rule-based with knowledge classification, Sohae.



Rule-Based, With Knowledge

Figure 13: Rule-based with knowledge classification, Yongbyon.

## 4.1.3. Nearest Neighbor Classification without Knowledge in eCognition

The first nearest neighbor classifications of each site used the no-knowledge feature space. Unlike the previous rule-based classifications, the remaining classifications were automated, so I recorded the time taken to compute each classification. The Sohae image classification with this algorithm took 46.025 minutes and resulted in 25 out of the 31 image objects being correctly classified as buildings. For the Yongbyon image, the algorithm took 1:45:32 hours to complete and resulted in 21 of the 27 image objects being correctly classified as buildings. The classification maps are as seen below:

# Nearest Neighbor, No Knowledge



Figure 14: NN no knowledge classification, Sohae.

0.9



Figure 15: NN no knowledge classification, Yongbyon.

## 4.1.4. Nearest Neighbor Classification with Knowledge in eCognition

These two nearest neighbor classifications used the with-knowledge feature space for classification. With the Sohae image, the algorithm took 23.08 minutes and resulted in 24 of 31 correctly classified buildings with 3 false positives. The Yongbyon image took the algorithm 34.06 minutes and resulted in 17 of 27 correctly classified buildings. The classification schemes of each for this algorithm are as follows:



# Nearest Neighbor, With Knowledge

Figure 16: NN with knowledge classification, Sohae.



# Nearest Neighbor, With Knowledge

4.1.5. Random Forest Classification without Knowledge in eCognition

The random forest classifications used two rules. To assess the amount of time the classifications required, I added the processing times of the two rules. The first two used the no-knowledge feature space for classification. For the Sohae image, training the classifier took 0.203 seconds, and applying the classifier took 1.703 seconds. This resulted in 25 of 31 buildings correctly classified. For the Yongbyon image, training the classifier took 0.613 seconds and applying the classifier took 7.851 seconds to classify the image. This resulted in 21 of 27 building image objects correctly classified. The classification maps for these algorithms are as follows:

# Random Forest, No Knowledge



Figure 18: RF no knowledge classification, Sohae.

0.6 1.2 Miles 0 0.15 0.3 0.9





2 Kilometers Figure 19: RF no knowledge classification, Yongbyon.

## 4.1.6. Random Forest Classification with Knowledge in eCognition

The next iteration of random forest classifications used the with-knowledge feature space for classification. For the Sohae image, training the classifier took 0.219 seconds and applying the classifier took 1.032 seconds, resulting in 25 of 31 correctly classified building image objects and 2 false positives. For the Yongbyon image, training the classifier took 0.765 seconds and 4.891 seconds to apply the classifier. This resulted in 17 of 27 correctly classified building image objects and 6 false positives. The classifications for the with-knowledge random forest classifications, and the accuracy assessment table for all classifications in this first analysis, are as follows:



# Random Forest, With Knowledge

Figure 20: RF with knowledge classification, Sohae.
Land Cover Classes Background Building s Buildun Vegetation Water 2 Kilometers 0 0.250.5 1 1.5

Random Forest, With Knowledge

Figure 21: RF with knowledge classification, Yongbyon.

Classification	Sohae, RB, NK	Sohae, RB, WK	Sohae, NN, NK	Sohae, NN, WK	Sohae, RF, NK	Sohae, RF, WK	YB, RB, NK	YB, RB, WK	<u>YB, NN, NK</u>	YB, NN, WK	YB, RF, NK	YB, RF, WK
TP	25	25	25	24	25	25	13	10	21	17	21	17
FP	0	0	1	3	0	2	1	1	8	3	13	6
<u>TN</u>	100	100	99	97	100	98	99	99	92	97	87	94
FN	6	6	6	7	6	6	14	17	6	10	6	10
<u>CA</u>	0.9542	0.9542	0.9466	0.9237	0.9542	0.9389	0.8819	0.8819	0.8583	0.8898	0.8976	0.874
<u>UA</u>	0.8065	0.8065	0.9615	0.7742	0.8065	0.8065	0.4815	0.4815	0.3704	0.7778	0.6296	0.6296
PA	1	1	0.9615	0.8889	1	0.9259	0.9286	0.9286	0.9091	0.7241	0.85	0.7391
<u>F</u>	0.4464	0.4464	0.4808	0.4138	0.4464	0.431	0.3171	0.3171	0.2632	0.375	0.3617	0.3399

Table 6: Knowledge Incorporation comparison accuracy assessment (above).



Table 7: Knowledge Incorporation Analysis Processing Times.

### 4.2. Software Comparison

#### 4.2.1. Random Forest Classification in eCognition

For the random forest software comparison, I used the Sohae image classification without feature space reduction. These results do not differ from those discovered in the knowledge incorporation analysis of this study. Accordingly, the results - including the classification map and the accuracy assessment – may be found in Section 5.1.5 of this study.

### 4.2.2. Random Forest Classification in R

The random forest classifications in *R* are divided into three models each with its own results. Prediction maps for each of the models were created based on their respective results. Model 1, which is the simple random forest model using pixel values, achieved an overall accuracy of 64 % with the validation dataset. The error matrix for Model 1 is as follows:

	Building	Buildup	Vege_mount	Water			
Building	3	1	0	0			
Buildup	0	5	4	3			
Vege_mount	0	1	8	0			
Water	0	1	0	2			
Table 8: Model 1 Accuracy Assessment							

The prediction map of the classifications for Model 1 may be seen below:



Figure 22: Prediction using Model 1 (z-value of pixel).

Model 2 achieved an overall accuracy of 93 % using the validation dataset. This accuracy was achieved after several iterations of the model to determine the best choice of *mtry* value. The most appropriate value was 10, which produced the accuracy. The error matrix for Model 2, are as follows:

	Building	Buildup	Vege_mount	Water
Building	4	0	0	0
Buildup	0	12	0	0
Vege_mount	0	1	7	1
Water Table 9: Model 2 Acc	0 curacy Assess	0 s <i>ment</i>	0	3

Accordingly, the prediction map of the classification from Model 2 is seen below:

#### **RF-distance Classified Image**



Figure 23: Prediction using RF Model 2 (RFsp with z).

For this model, I also calculated the variable importance, which is the proportion of the frequency of occurrences of the variable in the classification trees. The z value showed greater importance with distance-based variables maintaining high significance, as well. The following shows the variables with a variable significance greater than 2.

Variable	Importance
z	17.0727495
layer.21	8.8608239
layer.23	5.8005842
layer.5	5.1880802
layer.22	4.9068297
layer.24	3.9597840
layer.13	3.5027885
layer.27	2.4581336

Table 10: Model 2 variables with greater than 2 significance.

Lastly, Model 3 used only the distance variables for the random forest classification. It performed best on the validation dataset, producing an overall accuracy of 94 %. The error matrix for Model 3 is as follows:

	Building	Buildup	Vege_mount	Water				
Building	4	0	0	0				
Buildup	0	12	0	0				
Vege_mount	0	0	8	1				
Water	0	0	0	3				
Table 11: Model 3 Accuracy Assessment								

Below is the prediction map of the classification produced with Model 3:



RFsp Classified Image, no covariates

Figure 24: Prediction using RF Model 3 (RFsp, no z).

### 4.3. Spatial Resolution Comparison

The spatial resolution comparison analysis used the three classification algorithms used in the first analysis on NAIP imagery at 1m and 3m resolution. This analysis used only the knowledge-based classification. Each of these classifications was classifying the entire scene as opposed to extracting the building class. The bulk of the 100 points used for assessing accuracy were either water or vegetation from the reference image, with much lower frequencies of the remaining classes.

### 4.3.1. Rule-Based Classification 1m

The first rule-based classification I conducted on the 1m version of the facility image. This required a total of 9 rules for classification. The most accurate class of production was the "water"



Rule-Based, 1m

and "vegetation" classes, whereas there was confusion with 5 "buildup" points being classified as "building" and "vegetation." The higher accuracy in "vegetation" and "water" classes is likely due to the greater number of samples in these classes relative to the remaining classes. The classification is seen in Figure 25, and its respective contingency table is seen below:

	Reference Image							
	<u>RB_1</u>	<u>Water</u>	<u>Buildup</u>	<b>Building</b>	<b>Vegetation</b>	<u>Total</u>		
	<u>Water</u>	22	0	0	0	22		
	<u>Buildup</u>	0	0	0	0	0		
<b>Classification</b>	<b>Building</b>	0	1	0	0	1		
	<b>Vegetation</b>	0	4	0	73	77		
	<u>Total</u>	22	5	0	73	100		
Table 12. Dula ba	and 1m Annur	NOV ADDOOR	ont					

Table 12: Rule-based, 1m Accuracy Assessment

## 4.3.2. Rule-Based Classification 3m

The next rule-based classification was conducted on the 3m version of the image, for which I used the same 9 rules as the 1m rule-based classification. This classification achieved high accuracies in classifying water and vegetation but very low accuracy in built-up classification, which was confused entirely with vegetation. The classification map is seen in Figure 26, and its contingency table is as follows:





Figure 26: Rule-based classification, 3m

	Reference Image								
	<u>RB_3</u>	<u>Water</u>	<u>Buildup</u>	<b>Building</b>	<b>Vegetation</b>	<u>Total</u>			
	<u>Water</u>	22	0	0	0	22			
	<u>Buildup</u>	0	0	0	0	0			
<b>Classification</b>	<b>Building</b>	0	0	0	0	0			
	<b>Vegetation</b>	0	8	0	70	78			
	<u>Total</u>	22	8	0	70	100			

Table 13: Rule-based, 3m Accuracy Assessment

### 4.3.3. Nearest Neighbor Classification 1m

The nearest neighbor classification that I conducted on the 1m version of the facility image took approximately 45:02 minutes to complete. The water and vegetation classes achieved the highest accuracies with built-up and building classification experiencing low accuracies, with at least one correctly classified point in each class. The classification map for the nearest neighbor classification of the 1m image is seen in Figure 27, and its contingency table is as follows:



Nearest Neighbor, 1m

<u>Reference image</u>								
<u>NN_1</u>	<u>Water</u>	<u>Buildup</u>	<b>Building</b>	<b>Vegetation</b>	<u>Total</u>			
<u>Water</u>	22	0	0	1	23			
<u>Buildup</u>	0	1	0	0	1			
<u>Building</u>	0	5	1	3	9			
<b>Vegetation</b>	1	0	3	63	67			
<u>Total</u>	23	6	4	67	100			
	<u>NN_1</u> <u>Water</u> <u>Buildup</u> <u>Building</u> <u>Vegetation</u> <u>Total</u>	NN_1WaterWater22Buildup0Building0Vegetation1Total23	NN_1WaterBuildupWater220Buildup01Building05Vegetation10Total236	NN_1WaterBuildupBuildingWater2200Buildup010Building051Vegetation103Total2364	NN_1WaterBuildupBuildingVegetationWater22001Buildup0100Building0513Vegetation10363Total236467			

- **f** - ... - . - . . . . . . . . . .

Table 14: Nearest Neighbor 1m, Accuracy Assessment

### 4.3.4. Nearest Neighbor Classification 3m

For the second nearest neighbor classification for the 3m version of the image, the classification algorithm took 20:09 minutes to complete. Vegetation and water again classified all respective points correctly, and all 7 built-up points were incorrectly classified as vegetation. No points retained building classifications. The classification map is in Figure 28 and the contingency table is seen below:



### Nearest Neighbor, 3m

2m	
3111	

	Reference Image								
	<u>NN_3</u>	<u>Water</u>	<u>Buildup</u>	Building	<b>Vegetation</b>	<u>Total</u>			
	<u>Water</u>	22	0	0	0	22			
	<u>Buildup</u>	0	0	0	0	0			
<b>Classification</b>	<u>Building</u>	0	0	0	0	0			
	<b>Vegetation</b>	0	7	0	71	78			
	<u>Total</u>	22	7	0	71	100			

Table 15: Nearest Neighbor 3m, Accuracy Assessment

### 4.3.5. Random Forest Classification 1m

The random forest classification of the 1m resolution image took approximately 0.52 seconds to train the classifier with the given samples and an additional 6.03 seconds to apply the classification to the image. Water and building classes retained the highest accuracy of correctly classifying all of their respective points in the reference image. The built-up class experienced an increase in correctly identified objects, 3 of 6 correctly classified points, while vegetation experienced increased confusion amongst all other classes, though the majority of its points were correctly classified. The classification map is in Figure 29, and the contingency table for this random forest classification may be seen below:



# Random Forest, 1m

	Reference Image								
	<u>RF_1</u>	<u>Water</u>	<u>Buildup</u>	<b>Building</b>	<b>Vegetation</b>	<u>Total</u>			
	<u>Water</u>	22	0	0	3	25			
	<u>Buildup</u>	0	3	0	4	7			
<b>Classification</b>	<b>Building</b>	0	0	1	6	7			
	Vegetation	0	3	0	58	61			
	<u>Total</u>	22	6	1	71	100			

Table 16: Random Forest 1m, Accuracy Assessment

# 4.3.6. Random Forest Classification 3m

For the random forest classification of the 3m resolution image, training the classifier took 0.44 seconds and applying the classifier, 4.56 seconds. Water yielded the highest accuracy, similar to the other classifications, while built-up yielded the lowest, again, with all of its points classified as vegetation. Vegetation retained its high accuracy though with one point of confusion

with the water class. The classification map and contingency table for this classification are as follows, followed by a discussion of these results:



Random Forest, 3m

	Reference Image								
	<u>RF_3</u>	<u>Water</u>	<u>Buildup</u>	<b>Building</b>	<b>Vegetation</b>	<u>Total</u>			
	<u>Water</u>	22	0	0	1	23			
	<u>Buildup</u>	0	0	0	0	0			
<b>Classification</b>	<u>Building</u>	0	0	0	0	0			
	<b>Vegetation</b>	0	7	0	70	77			
	<u>Total</u>	22	7	0	71	100			
- hla d. T. Davadava	- <b>F</b> ( <b>O A</b> -	A							

Table 17: Random Forest, 3m Accuracy Assessment



Table 18: Spatial Resolution Analysis Classification Processing Times

#### 4.3.7. Overall Results

Each analysis produced a variety of different accuracies, but for this study in determining the best algorithm and parameters for this classification, the algorithm with the best accuracy from each section will be selected. Our results show that for the first analysis, the most accurate algorithm and feature space combination was both feature sets of the rule-based algorithm and the no knowledge random forest classification for the Sohae image. For the Yongbyon image, the highest accuracy was found in the random forest without knowledge classification. For the second analysis, Model 3 in *R* produced a slightly greater overall accuracy than the knowledge-incorporated random forest classification in *eCognition* of the Sohae image. Lastly, for the spatial resolution comparison, the 1m nearest neighbor and rule-based classifications produced the fewest amount of incorrect classifications and thereby the highest accuracy of the classification-resolution combinations.

### 5. Discussion

The goal of this research was to determine whether expert knowledge could be used to improve classification accuracy for missile test sites through a process of feature reduction. To this end, four analyses were conducted to evaluate the roles that classification method, spatial resolution, program, and the inclusion of expert knowledge affect the identification of buildings. This research was conducted in the face of some limitations and was restricted further by our delimitations.

### 5.1. Limitations

Our study contained several limitations to its full potential results. The most prominent limitation is the restriction to only two main study sites. Additionally, these study sites existed in the DPRK, one of the more restricted nations in the world. The location of these study sites proved problematic in the range of imagery available to public use. Accordingly, I were restricted to the 3m Planet Labs imagery, which did not appear to provide adequate resolution for the knowledge incorporation classification comparisons. Additionally, I had no direct means to access the sites for potential ground truthing, if necessary. These images are susceptible to tampering as well, which remains a minor yet critical limitation if the research is to be used for intelligence or military applications.

#### 5.2. Delimitations

I delimited this study in several ways, including our restriction of study sites to those in the DPRK for the initial analysis and to strictly the Michigan nuclear facility in the final analysis. The results of this study may accordingly only apply to these specific cases and not to other cases, despite the reusability between the DPRK and the Michigan workflows. Furthermore, I limited the analysis to GEOBIA instead of incorporating pixel-based analysis methods, which may have produced different results for detecting the buildings in these images. Lastly, I limited the software comparison to only eCognition and R and limited this analysis to the application of random forest classification in the respective software. Similarly, these results could be software and algorithm specific – expanding the options and combinations would likely produce different results.

#### 5.3. Analysis 1 – Knowledge Incorporation Comparison

The results from the knowledge incorporation did not seem to confirm our hypothesis that simpler automated classification of buildings would be more successful than using the complete image feature information. This feature space reduction seems to have confirmed or reinforced notions of timeliness at the cost of accuracy posed by previous work in the GEOBIA literature. Aside from the classifications themselves, the calculation of the algorithms took much less time with the knowledge-based feature space. This observation is consistent with that seen in Ma et al. (2015), which emphasizes decreased computing time with the reduced feature space but also a reduced accuracy. This latter point, however, is not confirmed with our analysis. Reducing the number of features to reflect the interpreter's analysis in the feature space did not seem to affect the classifications for the Sohae image as much as the Yongbyon image. For the Sohae image, the number of true positive remained the same for both with- and without-knowledge classifications, except one less true positive in the with-knowledge nearest neighbor classification.

The features used in the classification of this image seemed to reflect those best used to identify buildings already since the building features (such as green and brightness values) tended

to be of much greater value in buildings for this image than in the Yongbyon image. The most notable difference in the feature space reduction for this image is the slight increase in false positives, which is likely due to some confusion between spectrally similar buildings and built-up areas in the image. The Yongbyon image, on the other hand, experienced more dramatic changes by the reduced accuracy found in previous literature (Ma et al. 2015; Yu et al. 2006). All three classifications algorithms for the Yongbyon site classified fewer true positives when the feature space was reduced for knowledge incorporation. The number of false positives, however, either stayed the same or decreased, likely due in part to this increase in false negatives.

Based on the classification maps, both sites appeared to have greater misclassifications with the reduced feature space. The fewer features, as mentioned in Ma et al. (2015), permit the algorithms fewer variables to consider in classifications, thereby leading to likely greater misclassification. All images from both sites appear much more speckled, as built-up tends to be misclassified as vegetation in the Sohae image due to the spectral similarity of the mountainous terrain.

In the Yongbyon image, the notable confusion apparent in the maps is the increase in water features throughout the classification. The image contains some sporadic water features which retain similar spectral similarity to the vegetation in the image, which likely attributed to the sporadic water classifications throughout the vegetation class in the automated supervised algorithms. Despite this water misclassification, the knowledge incorporation refined the building features within the built-up features. The accuracy assessments do not appear to reflect this refinement, but the knowledge incorporation for both the nearest neighbor and random forest classifications of the Yongbyon image appear to better identify the buildings as opposed to creating a large cluster of building-classified image objects which inevitably contain the buildings of interest.

I chose to focus on only the "building" class for this analysis. This choice reflects that found in Marpu et al. (2008), which identified the most appropriate method of classifying an Iranian nuclear facility as creating two binary classes (class of interest and not class of interest) and classifying accordingly, which produced favorable results. Our classifications provide an indicator of the presence of buildings rather than identifying specific buildings and their extent, likely due to the image quality. Accordingly, a higher resolution image would provide more accurate results of the extent of the building class.

Additionally, while Castilla and Hay (2008) emphasize the requirement of multiple iterations of segmentation and classification to incorporate semantic meaning into image objects, the results of this study appear to support their findings. Providing a simple feature reduction appears to identify the locations of the classes of interest with a single iteration though multiple iterations would likely have provided better results and greater semantic meaningful image objects.

The dramatic difference between the Sohae and the Yongbyon images is likely due to the sheer size difference between the two and the heterogeneity of the Yongbyon image versus the Sohae image. The Sohae image contained 8,019 image objects and the Yongbyon image, 330,722 image objects, contributing to a large difference in the size of the image objects and the image's overall homogeneity difference. The image objects in the Sohae image, as a result, were more representative of visually-interpreted objects and contained relatively similar and more-distinguishable feature values. Accordingly, the user's accuracy for all of the Yongbyon

classifications regardless of feature space were significantly lower than the Sohae user's accuracy for any classification. Additionally, all other measured accuracies were slightly lower in the Yongbyon image than the Sohae image.

#### 5.4. Analysis 2 – Software Comparison

Before this comparison, it does not seem a comparison between the R and eCognition random forest applications had been previously undertaken. The comparison is further refined with the use of random forest classification for analysis. As mentioned in the previous discussion section, the results of the classification in eCognitinion yielded good accuracies with mostly true positives in classifying the buildings. For all three models constructed in *R*, all building classifications were correctly classified, though they represent a low number of the validation set. Each of the models in *R* also produced high overall accuracies. Though not the intent of this analysis, it suggests the increase in the amount of image data used for classification may complicate the classification and lead to misclassifications. The argument reinforces the strength of feature reduction for identifying certain classes in an image.

Of the three models in the *R* software analysis, Model 2 produced the most accurate map. It generalized a fair portion of the built-up area in the image, similar to the *eCognition* analysis. This confusion likely lies in the similarity in spectral values of the built-up area and the vegetation-mountain class. *eCognition* appeared to better represent this built-up class, though with some errors in classifying as "vegetation," since the former does not create large, properly distinguishable features as groups of the built-up area as does the *R* classification. The *R* analysis classified a large portion of the bottom-right of the image as "building," however, which the *eCognition* classification better identifies.

#### 5.5. Analysis 3 – Spatial Resolution Comparison

Per our predictions for the spatial resolution comparison of the Palisades Nuclear Facility image, the 1m resolution images provided better classification results across all three classification algorithms using the knowledge-based feature space. As was the issue in Maathuis (2003), where low-resolution imagery could not accurately classify individual landmines, the lower resolution imagery of the Palisades facility reduced accuracy. This observation supports the findings in the study above that higher resolution imagery – sub-meter resolution as designated by Maathuis (2003) – is needed to identify smaller features in spaceborne imagery properly.

The lower accuracies achieved in the 3m resolution image correlates to the lower accuracies reflected in our knowledge incorporation comparison analysis of this study. The 3m resolution data did not accurately classify "buildings" or any of the other used classes and the 1m resolution imagery with the given rulesets and classification algorithms. The validation points in the 3m resolution imagery did not represent the building class in any of the instances, likely due to the larger image object size likely placing it in an image object of vegetation or built-up classification. Even referring to the built-up class, the 3m resolution imagery did not correctly classify one of the built-up points from the reference image. Rather, these were all classified as vegetation, again likely due to the over-segmenting of the image objects representing vegetation into built-up areas being improperly classified as vegetation due to their spatial proximity to the vegetation classes. Shadows being classified as vegetation is not an expected outcome from our hypotheses; this confusion may be due to the darkness of the vegetation in certain parts of the image. Some

samples used for vegetation may have extracted these dark-appearing spectral values and did not have any other class appropriate for classifying the shadows than "vegetation." Creating a "shadows" class may mitigate this confusion.

Though the 3m resolution imagery did not yield high accuracy, the 1m resolution imagery was not perfect, and yielded higher, though not high, accuracy. The image objects in the 1m resolution image merged to represent buildings due largely to their high brightness value in this image. The finer imager objects produced allowed an easier merging process that permitted the merging of image objects to represent urban and artificial features correctly. The 1m imagery better defined the different features and prevented much of the over-segmentation present in the 3m resolution image due to finer, more-definable features. Of the three classifications used on the 1m resolution imagery, the rule-based classification remained the only algorithm not to identify building features at any of the random points used for verification. The other algorithms identified buildings in one or more instances at the 1m resolution level but produced mostly misclassifications as vegetation class. This confusion is likely due to the spatial proximity of the classes and the spectral similarity of vegetation and some of the shadows present near buildings.

The sample differences between the images may have contributed to the different accuracies in classification. As seen in the Methods section of this study, the samples for the 1m resolution imagery were much smaller and did not account for as much space as the 3m resolution samples. In attempting to transfer the TTA mask created from this 1m sampling, no accurate transfer was possible without large image objects being used as samples in the 3m resolution image. Accordingly, the 3m resolution samples likely accounted for largely different spectral values than the more refined 1m resolution samples. This likely contributed to the error seen in classifying buildings, in particular. The misclassification of the vegetation image objects in the 1m resolution imagery may result from the smaller image objects accounting for a more limited threshold of values and the increased number of image objects in general due to greater heterogeneity in the finer resolution image.

Both of the resolutions accurately classified non-urban features. The 1m random forest classification produced the lowest accuracy for classifying vegetation, however, with confusion between water and built-up classes. The confusion may lie in the spectral similarity of shadows, darker water features, and darker vegetation, a problem which may potentially be solved with the addition of a shadows class. Water features were classified correctly at all identified points, except one point in the finer resolution nearest neighbor classification, likely due to oversegmenting or spectral similarity of very close image objects to water features.

Even though the 1m resolution image classification produced more favorable results than the 3m resolution image, the algorithms took a longer duration to complete. For the rule-based classifications, I used the same 9 rules to classify the images, while the nearest neighbor took almost double the amount of time for the 1m resolution image than it needed for the 3m resolution image. The random forest classification took about one-second longer for the 1m resolution image classification. This difference in time reflects the greater amount of information present in the 1m resolution image objects to classify than di the 3m resolution image. The larger image objects in the 3m resolution image objects in the 1m resolution image.

Based on the classification maps, the lower resolution imagery appears to provide cleaner classifications than the 1m resolution imagery. This observation may likely be attributed to the far greater number of image objects in the finer resolution imagery. Additionally, the finer resolution imagery contains an additional "residential" class, which could not be classified in the 3m resolution imagery due to large spectral confusion in classifying the entire scene as this class. As seen in the 1m resolution classifications, the "residential" class appears sporadically confused with "vegetation" and "industrial building" classes. To a degree, these may display accurately, since residential buildings were represented by small image objects throughout the image; however, the similarity between the residential image objects and the surrounding "vegetation," "industrial," and "built-up" classes. The 3m classifications did not produce as much apparent noise as the 1m resolution classifications. Accordingly, they appear to have identified specific buildings better than the 1m classifications, which produce a fair amount of noise for the "building" and "built-up" classes, again likely due to the similarity between the "built-up" and "building" classes spectrally, and the increased number of image objects. These 3m classifications, however, did produce unclassified image objects in the final classification. The rule-based 1m resolution classification appears to produce the most realistic classification, including the residential buildings divided, whereas the random forest from the same resolution produces a noise-filled classification with several apparent misclassifications.

#### 5.6. Developments on Present Research

This research builds on and solidifies some of the previous research regarding knowledge incorporation into GEOBIA applications and applications observing the DPRK and military applications with remote sensing analysis. The research extends previous research regarding the

DPRK from a geographic perspective, as designated by Shim 2014 (Shim 2014). It additionally details the seldom-analyzed missile testing facility components specifically, which often are overshadowed by the missile development and the missile program in the DPRK itself. Regarding remote sensing for strategic operations, this research contributes by confirming the need for high-resolution imagery to detect small features in images, as observed by Maathuis (Maathuis 2003). These analyses further show that conducting remote sensing studies reflecting strategic operations may be conducted to a lesser degree with publicly available data for little to no cost to the analyst.

There appears to be a large amount of research towards the development of contextual, knowledge-incorporated image classification in the realm of GEOBIA. The research produced in this study, particularly our first analysis, builds on the assertion by MacFaden and O'Niel-Dunne (2015) that different rulesets will be necessary to properly classify different sites with contextual information. The comparison of different classification methods is similar to Belgiu and Dragut (2016), who compare and supervised classification methods for classification accuracy. Additionally, the methods in this study may motivate to further integrate eCognition with the R computer software, which has not been directly compared in previous research as far as I am aware.

### 6. Conclusion

In our research study, I aimed to observe the best approach of incorporating expert knowledge into classifications of DPRK missile testing facilities. The study compared three different perspectives and variables associated with image analysis. Accordingly, I divided the study into three analyses, one for each of these factors: knowledge incorporation based on feature reduction, software for classification, and spatial resolution of the image for classification. Though our specific analyses did not produce outstanding results, it still narrowed the algorithms in the study to a most accurate combination across the three analyses, based on the most accurate parameters from each analysis. In several cases, I conclude that the different classification methods and combinations of analyses results could be used with some success in different scenarios. By the results provided above, however, I conclude that the best results provided from our analyses based on the comparison is a knowledge-based, random forest classification using R and 1-meter spatial resolution. Different applications of the research results would require different combinations of the best-resulting variables. Overall, the intent of the study remained to see to what extent I could conduct these analyses using publicly-available data to perform analyses for intelligence-like applications in a politically isolated nation of the world critical to government intelligence interests.

Several different variables could be accounted for in this study to improve the results for future refinement and research. More research could be done to further the results produced in this study. The use of strictly 3m resolution data for the DPRK classifications could be improved to sub-meter resolutions for likely better analysis and a more accurate representation of each of the algorithms and their potentials. Since I restricted the study to only DPRK sites, future research could expand this analysis to similarly rogue nations with nuclear and missile programs, namely the IRI, which has been the greater focus of many similar studies. These less-accessible regions appear to be one of the only blockades toward achieving entirely public information for similar studies.

According to these future research suggestions to improve the motives of this study, some additional factors could be tested to develop and discover the best-fitting combination of variables for this political application of image classification. A study of different combinations of the results of this study may provide a more diversified analysis of the accuracies of different combinations of data, classification algorithms, knowledge incorporation, and software. For instance, a future study could take the most accurate approaches found in this study and test the three variables together for accuracy. It may add to the study with different combinations to compare the accuracies. Performing this combination again on the DPRK sites used in this study will also provide a more in-depth analysis of the research area and further refine the results of this research.

While this study focused on GEOBIA applications primarily, it may be an interesting analysis to compare the results of the different algorithms for building identification between pixel-based methods and GEOBIA methods. The introduction of pixel-based analysis would furthermore introduce new software to the software comparison of this study, including ERDAS Imagine and ArcGIS programs, to name a few. This software could be added to the software comparison to expand on this particular analysis. Additionally, higher and lower resolution images could be used for classification to determine the effects of increasing and decreasing the resolution for each of the classification algorithms. I would also like to compare highly quantitative methods developed

in GEOBIA literature to those that are highly qualitative methods to determine which approach provides the most accurate classifications. All of these analyses could be used on the DPRK missile sites to advance the research done in this study in the interests of supporting policymaking and government decision-making.

I hope that the results of this study motivate to continue geographic research in the DPRK. Little research has been conducted in the region that is available to the public. While some may believe it impossible to conduct geographic research of the DPRK without classified governmentlevel imagery, this study shows that such analysis may be conducted with publicly available assets and imagery for little cost to the individual. The DPRK missile program remains a highly significant issue in global affairs, particularly in its threats towards the US and its allies historically and recently. With this issue becoming publicly available, it is critical that the public view the program on the ground to truly understand it and the DPRK capabilities so as to avoid inappropriate political moves or gestures. Using the ever-advancing field of remote sensing and GEOBIA and results of this study, the public may be able to make further developments and refinements to develop an accurate representation of highlighting the buildings in DPRK missile testing facilities for better public perception of the program and its developments as they occur using entirely open-source data.

APPENDICES

### **APPENDIX A: Abbreviations**

- CONAN: Contextual Analysis
- **<u>CTB</u>**: Comprehensive Test Ban Treaty
- **<u>DPRK</u>**: The Democratic People's Republic of Korea (The DPRK)
- FMLE: Fuzzy Maximum Likelihood Estimation
- GEOBIA: Geographic Object-Based Image Analysis
- IC: Intelligence Community
- ICBM: Intercontinental Ballistic Missile
- ID: Identification
- IRI: Islamic Republic of Iran
- MAD: Multivariate Alteration Detection
- **<u>NPT</u>**: Nuclear Nonproliferation Treaty
- **NTI**: Nuclear Threat Initiative
- **<u>OWL</u>**: Ontology Web Language
- **<u>ROK</u>**: The Republic of Korea (South Korea)
- **<u>SEaTH</u>**: Separability and Threshold
- **<u>SLBM</u>**: Submarine Launched Ballistic Missile
- US: The United States of America
- **<u>USGS</u>**: United States Geological Survey
- **USSR**: United Soviet Socialist Republics (Soviet Union)
- **<u>WMD</u>**: Weapons of Mass Destruction

#### **APPENDIX B: Sohae Ruleset**

⊟- ■ sohae · segmentation -- 50 [shape:0.1 compct:0.5] creating 'Level 1' with Brightness = 0 at Level 1: merge region with Number of pixels >= 9200 at Level 1: merge region B- classification ⊨ • water tunclassified with ndvi <= -0.5 at Level 1: water water at Level 1: merge region tunclassified with Max. diff. >= 1 at Level 1: water 1 water with Rel. border to unclassified = 1 at Level 1: unclassified 🔥 unclassified with Rel. border to water >= 0.6 and Rel. border to buildup >= 0 at Level 1: water 1. with Brightness = 0 at Level 1: background E- • buildings 1 with Brightness >= 4500 at Level 1: building E- • buildup 1 unclassified with Standard deviation red >= 120 at Level 1: buildup 🔥 unclassified with Brightness >= 4000 at Level 1: buildup 1 unclassified with Rel. border to buildup >= 0.4 at Level 1: buildup www.buildup at Level 1: merge region . vegetation\_mountain - 🔼 unclassified at Level 1: vege\_mountain in- reset - supervised 🔥 at Level 2: background, building, buildup, vege\_mountain, water - . RF class 50 [shape:0.1 compct::0.5] creating 'Level 3' 1. 0.203 at Level 3: classifier: train random trees using Brightness, Max. diff., blue, green, nir, red, blue, green, nir, red, Area, Border length, Rel. border to image border, Volume, Asymmetry, Border index, Compactness, Density, Elliptic fit, Main direct -1. 01.703 at Level 3: classifier: apply Mat Level 3: remove classification - • knowledge - 50 [shape:0.1 compct::0.5] creating 'Level 4' www at Level 4: merge region E- th 1 with Brightness = 0 at Level 4: background 1 with Mean nir > 0 and Mean nir <= 2100 at Level 4: water - the with Mean green > 0 and Mean green <= 4200 at Level 4 vege mountain 1 with Brightness <= 3300 at Level 4 vege\_mountain tunclassified at Level 4: buildup 1 buildup with Mean green <= 10000 and Mean green >= 5000 at Level 4 building i nn 10.016 at Level 6: copy creating 'Level 5' below No at Level 5: remove classification 批 at Level 5: background, building, buildup, vege\_mountain, water 8- • ff TT at Level 5: copy creating 'Level 6' below Hat Level & remove classification 🙏 0.219 🛛 at Level 6: classifier: train random trees using Brightness, green, nir, Area, Number of pixels, Asymmetry, Elliptic fit, Rectangular fit, Roundness 1.032 at Level 6: classifier: apply

#### **APPENDIX C: Yongbyon Ruleset**

```
    Yongbyon

  🐱 delete 'New Level'
  ₩ delete 'New Level1'
  My at New Level: remove classification
segmentation
    -->= 25 [shape:0.1 compct::0.5] creating 'Level1_1'
    at Level1_1: merge region
- classification
    - 1 with ndvi >= 0.05 at Level1_1: vegetation
      L unclassified with Mean NIR <= 3050 at Level1_1: water
      1 with Mean Blue = 0 at Level1_1: background
      1 unclassified with Max. diff. > 0 and Max. diff. <= 0.45 at Level1_1: buildup
      L unclassified with Brightness >= 4300 at Level1_1: buildup
     Lunclassified at Level1_1: vegetation
      1. buildup with Mean NIR <= 4000 and Mean NIR >= 1 at Level1_1: buildings
     1 buildup with Brightness >= 6600 at Level1_1: buildings
⊡-• nn
     T at Level1_1: copy creating 'Level1_nn' below
     Lat Level1_nn: background, buildings, buildup, vegetation, water
à.• f
      37 background at Level1_nn: copy creating 'Level1_rf' below
    at Level1 rf: remove classification
      批 at Level1_rf: classifier: train random trees using Blue, Brightness, Green, Layer 1, Layer 2, Layer 3, Layer 4, Max. diff., NIR, Red, Blue, Green, Layer 1, Layer 2, Layer 3, Layer 4, NIR, Red, Number of pixels, Asymmetry, Border index, Compactness, Density,
    A Level1_rf: classifier: apply
- • knowledge based
   . unsupervised
         TT at Level1_rf: copy creating 'Level2_1' below
        → at Level2 1: remove classification
         1 with Mean NIR > 0 and Mean NIR <= 3000 at Level2_1: water
         L unclassified with Mean Green > 0 and Mean Green <= 5000 at Level2 1: vegetation
         1: buildup and Mean Green >= 5000 and Mean Green <= 11000 at Level2_1: buildup
         1 buildup with Brightness >= 6500 at Level2_1: buildings
         1 buildup with Mean NIR <= 4000 and Mean NIR >= 3000 at Level2_1: buildings
   i nn
         T at Level2_rf: copy creating 'Level2_nn1' below
         1 at Level2_nn1: background, buildings, buildup, vegetation, water
   6 . f
        - TT at Level2_nn1: copy creating 'Level2_f1' below
         🤱 at Level2_rf1: classifier: train random trees using Green, Area, Border length, Number of pixels, Roundness, Rectangular fit, Asymmetry, Brightness, NIR
         1 at Level2_rf1: classifier: apply
```

## **APPENDIX D: Palisades Nuclear Energy Facility Ruleset**



REFERENCES

#### REFERENCES

38 North. 2018. 38 North. https://www.38north.org/.

- Agnew, John, Thomas W. Gillespie, Jorge Gonzalez, and Brian Min. 2008. "Baghdad Nights: Evaluating the US Military 'Surge' Using Nighttime Light Signatures." *Environment and Planning* 40: 2285-2295.
- Albright, David, and Paul Brannan. 2007. *The North Korean Plutonium Stock, February 2007.* Institute for Science and International Security.
- Arvor, Damien, Laurent Durieux, Samuel Andrés, and Marie-Angélique Laporte. 2013. "Advances in Geographic Object-Based Image Analysis with Ontologies: A Review of Main Contributions and Limitations from a Remote Sensing Perspective." ISPRS Journal of Photogrammetry and Remote Sensing 125-137.
- Baatz, M, C Hoffmann, and G Willhauck. 2008. "Processing from Object-Based to Object-Oriented Image Analysis." In Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications, by T Blaschke, S Lang and G J Hay, 29-42. Berlin: Springer-Verlag Berlin Heidelberg.
- Baatz, M., and A. Schäpe. 2010. *Multiresolution Segmentation: An Optimization Approach for High Quality Multi-Scale Image Segmentation.* http://www.agit.at/papers/2000/baatz\_FP\_12.pdf.
- Bazeley, Pat, and Kristi Jackson. 2013. *Qualitative Data Analysis with NVIVO*. 2nd. Los Angeles: SAGE Publications Inc.
- Beck, Richard A. 2003. "Remote Sensing and GIS as Counterterrorism Tools in the Afghanistan War: A Case Study of the Zhawar Kili Region." *The Professional Geographer* 170-179.
- Belgiu, Mariana, and Lucian Dragut. 2016. "Random Forest in Remote Sensing: A Review of Applications and Future Directions." *ISPRS Journal of Photogrammetry and Remote Sensing* 24-31.
- Belgiu, Mariana, Barbara Hofer, and Peter Hofmann. 2014. "Coupling Formalized Knowledge Bases with Object-Based Image Analysis." *Remote Sensing Letters* 530-538.
- Benz, Ursula C, Peter Hofmann, Gregor Willhauck, Iris Lingenfelder, and Markus Heynen. 2004. "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information." *ISPRS Journal of Photogrammetry and Remote Sensing* 239-258.
- Bhaskaran, Sunil, Shanka Paramananda, and Maria Ramnarayan. 2010. "Per-pixel and Object-Oriented Classification Methods for Mapping Urban Features Using Ikonos Satellite Data." *Applied Geography* 650-665.

- Bianchetti, Raechel, and Alan MacEachren. 2015. "Cognitive Themes Emerging from Air Photo Interpretation Texts Published to 1960." *ISPRS International Journal of Geo-Information* 551-571.
- Blaschke, Thomas, Geoffrey J Hay, Maggi Kelly, Stefan Lang, Peter Hofmann, Elisabeth Addink, Raul Queiroz Feitosa, et al. 2014. "Geographic Object-Based Image Analysis – Towards a new paradigm." *ISPRS Journal of Photogrammetry and Remote Sensing* 180-191.
- Breiman, Leo. 2001. "Random Forests." *Machine Learning* 5-32.
- Broad, WJ, D Jehl, DE Sanger, and T Shanker. 2005. "North Korea Nuclear Goals: Case of Mixed Signals." *New York Times.*
- Castilla, G, and G J Hay. 2008. "Image Objects and Geographic Objects." In *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, by T Blascke, S Lang and G J Hay, 91-110. Berlin: Springer-Verlag Berlin Heidelberg.
- Chen, W, X. Li, Y. Wang, and G. Liu, S. Chen. 2014. "Forested Landslide Detection Using LIDAR Data and the Random Forest Algorithm: A Case Study of the Three Gorges, China." *Remote Sensing of Environment* 291-301.
- Chung, Samman. 2016. "North Korea's Nuclear Threats and Counter-Strategies." *The Journal of East Asian Affairs* 83-131.
- Cloud, John G, and Keith C Clarke. 1999. "Through a Shutter Darkly: The Tangled Relationship Between Civilian, Military, and Intelligence Remote Sensing in the Early U.S. Space Program." In Secrecy and Knowledge Production, by Judith Reppy, 36-56. Cornell University Peace Studies Program.
- Diamond, John M. 2001. "Re-Examining Problems and Prospects in U.S. Imagery Intelligence." International Journal of Intelligence and Counterintelligence 1-24.
- Dragut, Lucian, and Thomas Blaschke. 2006. "Automated Classification of Landform Elements Using Object-Based Image Analysis." *Geomorphology* 330-344.
- Glade, David. 2000. Unmanned Aerial Vehicles: Implications for Military Operations. Unclassified Military Report, Maxwell AFB, AL: Air University Press.
- Gu, H. Y., H. T. Li, L. Yan, and X. J. Lu. 2015. "A Framework for Geographic Object-Based Image Analysis (GEOBIA) Based on Geographic Ontology." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 21-23.
- Gupta, Vipin, and Frank Pabian. 1997. "Investigating the Allegations of Indian Nuclear Test Preparations in the Rajasthen Desert." *Science & Global Security* 101-188.
- Hay, G.J., and G. Castilla. 2008. "Geographic Object-Based Image Analysis (GEOBIA): A New Name for a New Discipline." In *Object-Based Image Analysis: Spatila Concepts for Knowledge-Driven Remote Sensing Applications*, by Th. Blaschke, S. Lang and G.J. (Eds.) Hay, 74-89. Berlin, Heidelberg: Springer.

- Hitchings, Sean. 2003. "Policy Assessment of the Impacts of Remote-Sensing Technology." *Space Policy* 119-125.
- Hsieh, Hsiu-Fang, and Sarah E Shannon. 2005. "Three Approaches to Qualitative Content Analysis." *Qualitative Health Research* 1277-1288.
- Johnson, Brian, and Zhixiao Xie. 2013. "Classifying a High Resolution Image of an Urban Area using Super-object Information." *ISPRS Journal of Photogrammetry and Remote Sensing* 40-49.
- Kim, Sung Chull, and Michael D Cohen. 2017. North Korea and Nuclear Weapons: Entering the New Era of Deterrence. Washington DC: Georgetown University Press.
- Kim, Won-Young, and Paul G. Richards. 2007. "North Korean Nuclear Test: Seismic Discrimination Low Yield." *Eos* 158-161.
- Kit, Oleksandr, and Matthias Lüdeke. 2013. "Automated Detection of Slum Area Change in Hyderbad, India using Multitemporal Satellite Imagery." *ISPRS Journal of Photogrammetry and Remote Sensing* 130-137.
- Krtalic, A. 2016. "Analysis of the Segmented Features of Indicator of Mine Presence." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 12-19.
- Lexical Computing CZ. 2019. Sketch Engine. Mikulov, CZ.
- Liedtke, C.-E, J Buckner, O Grau, S Growe, and R Tonjes. 1997. "AIDA: A System for the Knowledge Based Interpretation of Remote Sensing Data." *Third International Airborne Remote Sensing Conference and Exhibition.* Copenhagen.
- Ma, Lei, Liang Cheng, Manchun Li, Yongxue Liu, and Xiaoxue Ma. 2015. "Training Set Size, Scale, and Features in Geographic Object-Based Image Analysis of Very High Resolution Unmanned Aerial Vehicle Imagery." *ISPRS Journal of Photogrammetry and Remote Sensing* 14-27.
- Maathuis, B. H.P. 2003. "Remote Sensing Based Detection of Minefields." *Geocarto International* 51-60.
- MacFaden, Sean, and Jarlath O'Neil-Dunne. 2015. "A Tool for the Automated Detection of Damaged Transportation Infrastructure." *ASPRS Annual Conference*. Tampa.
- Marpu, P R, I Niemeyer, S Nussbaum, and R Gloaguen. 2008. "A Procedure for Automatic Object-Based Classification." In *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, by T Blaschke, S Lang and G J Hay, 169-184. Berlin: Springer-Verlag Berlin Heidelberg.
- Maxwell, A.E., T.A. Warner, M.P. Strager, J.F. Conley, and A.L. Sharp. 2015. "Assessing Machine-Learning Algorithms and Image- and Lidar-derived Variables for GEOBIA Classification Mining and Mine Reclamation." *International Journal of Remote Sensing* 954-978.
- McTavish, Donald G, and Ellen B Pirro. 1990. "Contextual Content Analysis." *Quality and Quantity* 245-265.
- Millard, K., and M. Richardson. 2015. "On the Importance of Training Data Sample Selection in Random Forest Image Classification: A Case Study in Peatland Ecosystem Mapping." *Remote Sensing* 8489.
- Ming, Dongping, Jonathan Li, Junyi Wang, and Min Zhang. 2015. "Scale Parameter Selection by Spatial Statistics for GEOBIA: Using Mean-Shift Based Multi-Scale Segmentation as an Example." *ISPRS Journal of Photogrammetry and Remote Sensing* 28-41.
- Niemeyer, I, P R Marpu, and S Nussbaum. 2008. "Change Detection using Object Features." In *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, by T Blaschke, S Lang and G J Hay, 185-202. Berlin: Springer-Verlag Berlin Heidelberg.
- Niemeyer, I., and S. Nussbaum. n.d. "Automation of Change Detection Procedures for Nuclear Safeguards-Related Monitoring Purposes." Global Monitoring for Security and Stability Network of Excellence.
- Nuclear Threat Inititative. 2018. North Korea. http://www.nti.org/learn/countries/north-korea/facilities/.
- Nussbaum, S., I. Niemeyer, and M. J. Canty. 2006. "SEATH A New Tool for Automated Feature Extraction in the Context of Object-Based Image Analysis."
- O'Neil-Dunne, Jarlath P.M., Sean MacFaden, and Keith C Pelletier. 2011. "Incorporating Contextual Information into Object-Based Image Analysis Workflows." *ASPRS 2011 Annual Conference*. Milwaukee.
- Ozeki, Masaru, and Kosuke Heki. 2010. "Ionospheric Holes made by Ballistic Missiles from North Korea Detected with a Japanese Dense GPS Array." *Journal of Geophysical Research* 115.
- Pal, M., and P.M. Mather. 2005. "Support Vector Machines for Classification in Remote Sensing." International Journal of Remote Sensing 1007-1011.
- Perkins, Chris, and Martin Dodge. 2009. "Satellite Imagery and the Spectacle of Secret Spaces." *Geoforum* 546-560.
- Planet Labs. 2019. Planet Labs Imagery & Archive. Mountain View, CA.
- Pollack, Jonathan D. 2003. *The United States, North Korea, and the End of the Agreed Framework.* US Naval War College.
- Postol, Theodore, and Markus Schiller. 2016. "The North Korean Missile Program." *Korea Observer* 751-805.
- Radoux, J, P Bogaert, D Fasbender, and P Defourney. 2011. "Thematic Accuracy Assessment of Geographic Object-Based Image Classification." *International Journal of Geographic Information Science* 895-911.

- Sachdov, A.V. 2000. "North Korea's Missile Programme: A Matter of Concern." *Strategic Analysis* 1695-1707.
- Saldaña, Johnny. 2013. *The Coding Manual for Qualitative Researchers.* 2nd. Los Angeles: SAGE Publications Inc.
- Samad, Tariq, John S. Bay, and Datta Godbole. 2007. "Network-Centric Systems for Military Operations in Urban Terrain: The Role of UAVs." *Institute of Electrical and Electronics Engineers.*
- Satyanarayana, P, and S. Yogendran. 2013. "Military Applications of GIS." *IIC Technologies Private Limited, Hyderabad.*
- Schlittenhardt, J, M Canty, and I Grunberg. 2010. "Satellite Earth Observations Support CTBT Monitoring: A Case Study of the Nuclear Test in North Korea of Oct. 9, 2006 and Comparison with Seismic Results." *Pure and Applied Geophyics* 601-618.
- Shim, David. 2014. "Remote Sensing Place: Satellite Images as Visual Spatial Imaginaries." *Geoforum* 51: 152-160.
- Shippert, Peg. n.d. Introduction to Hyperspectral Image Analysis. Research Systems, Inc.
- Squassoni, Sharon A. 2005. North Korea Nuclear Weapons: How Soon an Arsenal? CRS Report for Congress, Congressional Research Service.
- Stanford Center for Biomedical Informatics Research. 2016. Protégé. Stanford, CA.
- Taubenböck, H., T. Esch, M. Wurm, A. Roth, and S. Dech. 2010. "Object-Based Feature Extraction Using High Spatial Resolution Satellite Data of Urban Areas." *Journal of Spatial Science* 117-132.
- The R Foundation. 2019. *R version 3.5.3*. Murray Hill, NJ.
- Torres-Sánchez, J., F. López-Granados, and J.M. Peña. 2015. "An Automatic Object-Based Method for Optimal Thresholding in UAV Images: Application for Vegetation Detection in Herbaceous Crops." *Computers and Electronics in Agriculture* 43-52.
- Trimble. 2019. eCognition Developer 9. Sunnyvale, CA.
- Tuxen, K., and M. Kelly. 2008. "Multi-Scale Functional Mapping of Tidal Marsh Vegetation Using Object-Based Image Analysis." In Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications, by Th. Blaschke, S. Lang and G.J. Hay, 415-442. Berlin Heidelberg: Springer-Verlag.
- United States Department of Agriculture. 2015. *National Agriculture Imagery Program.* Salt Lake City, UT.

United States Geological Survey. 2019. *Earth Explorer*. Washington DC.

- Wang, H., Y. Zhao, R. Pu, and Z. Zhang. 2015. "Mapping Robinia Pseudoacacia Forest Health Conditions by Using Combined Spectral, Spatial, and Texural Information Extracted from IKONOS Imagery and Random Forest Classifier." *Remote Sensing* 9020.
- Watts, A.C., L.N. Kobziar, and H.F. Percival. 2009. "Unmanned Systems for Wildland Fire Monitoring and Research." *24th Tall Timbers Fire Ecology Conference*. Tallahasee. 86-90.
- WC3 Web Ontology Working Group. 2004. OWL2.
- Weinberger, Kilian Q, John Blitzer, and Lawrence K Saul. 2006. "Distance Metric Learning for Large Margin." Advances in Neural Information Processing Systems.
- Wharton, Stephen W. 1982. "A Contextual Classification Method for Recognizing Land Use Patterns in High Resolution Remotely Sensed Data." *Pattern Recognition* 317-324.
- Witmer, Frank D.W. 2015. "Remote Sensing of VioeInt Conflict: Eye from Above." *International Journal of Remote Sensing* 2326-2352.
- Yu, Qian, Peng Gong, Nick Clinton, Greg Biging, Maggi Kelly, and Dave Schrikauer. 2006. "Object-Based Detailed Vegetation Classification with Airborne High Spatial Resolution Remote Sensing Imagery." *Photogrammetric Engineering & Remote Sensing* 799-811.
- Yue, Peng, Liping Di, Yaxing Wei, and Weiguo Han. 2013. "Intelligent Services for Discovery of Complex Geospatial Features from Remote Sensing Imagery." *ISPRS Journal of Photogrammetry and Remote Sensing* 151-164.