REMOTE SENSING AND COMPUTER VISION ALGORITHMS AT SCALE: DEFENSE AND HUMANITARIAN USES

A Thesis submitted to the faculty of

San Francisco State University

In partial fulfillment of

the requirements for

the Degree

Master of Science

In

Geographic Information Science

by

Logan J. Wenzler

San Francisco, California

May 2019

CERTIFICATION OF APPROVAL

I certify that I have read Remote Sensing and Computer Vision Algorithms at Scale: Defense and Humanitarian Uses by Logan J. Wenzler, and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirement for the degree Master of Science in Geographic Information Science at San Francisco State University.

Leonhard Blesius, Ph.D.
Associate Professor

Jerry Davis, Ph.D.
Professor

REMOTE SENSING AND COMPUTER VISION ALGORITHMS AT SCALE: DEFENSE AND HUMANITARIAN USES

Logan J. Wenzler

San Francisco, California

2019

The objective of this thesis is to understand how cloud computing and artificial intelligence can be applied to vast amounts of remotely sensed data to better understand macro-level trends for humanitarian and defense issues. Computer vision algorithms and data were provided by Orbital Insight, Inc., a geospatial analytics company based in Palo Alto, CA.

Specific projects were curated, data was acquired, and analysis was applied to three use cases: "Patterns of life for The Battle of Marawi", "Indications and Warnings using multi-class aircraft detections", and "Camp Fire land cover analysis". The use cases show how with imagery ingestion pipelines, cloud computing, and computer vision algorithms, a massive quantity of data can be analyzed in a relatively short amount of time. Without these workflows and new technologies, analysis of large amounts of data would prove to be less efficient and resource heavy. The events show the benefits users of spatial data would have to gain a better understanding of humanitarian and defense issues. Algorithms used to derive insights from thousands of imagery scenes consisted of a car detection algorithm, multi-class aircraft algorithm, and a land cover classification algorithm. Additionally, the thesis briefly explores the use of geolocation data to supplement computer-vision algorithm data. The thesis shows on a high level, through examples, how users could use the technologies to analyze data more efficiently. This analysis can be incorporated into high-level humanitarian or defense decisions. Future work regarding this field should seek to evaluate the algorithm performance on a more granular level. Researchers should also build different algorithms on open-source imagery to allow for more users to benefit from the efficiency computer vision provides.

I certify that the Abstract is a correct representation of the content of this thesis.

L. [signature]
Chair, Thesis Committee

21. May 2019
Date

ACKNOWLEDGEMENTS

First, thank you to Dr. Leonhard Blesius for being a flexible instructor and advisor. Without his patience, balancing work and school surely would have driven me mad. Also, thanks to Dr. Jerry Davis, Dr. Ellen Hines, and Dr. XiaoHang Liu for all the GIS and Remote Sensing instruction that led to my current employment. Additionally, thank you to Dr. Sara Baguskas, Dr. Jennifer Blecha, and Dr. Nancy Wilkinson for thought-provoking class discussions. Thank you to all the students who helped shape a positive graduate experience. Lastly, thank you to Orbital Insight for providing me access to cutting edge technology and data which allowed me to create unique use cases for this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Most classification and object detection methods are applied to a few images at a time. Artificial intelligence (AI) and cloud computing, combined with an increased availability of remotely sensed images, can increase the efficiency and effectiveness of analyzing vast amounts of imagery. By understanding macro-level trends such as disposition of military aircraft over hundreds of airfields, users understand patterns to help guide policy level or mission-oriented decisions. This is also true for defense and humanitarian issues. To show the value of rapid imagery ingestion pipelines, cloud computing, and computer vision (CV) algorithms, different real-world examples will be evaluated.

First, a brief overview of AI, CV, and machine learning (ML) will be explored to provide background on what these terms mean regarding remotely sensed satellite imagery. Advances in cloud computing and the use of graphics processing units (GPUs) have significantly advanced the ways AI/ML/CV analyze imagery and other forms of data (Jermain et al., 2016). The advances surrounding AI/ML/CV and cloud computing/processing power would not be possible without the actual data. In this case, the data is commercially available satellite imagery. New imagery providers and the reduced price of satellites and imagery have created an increase in imagery over the last decade (Chisuwo, 2018). Chapter 1 explores the nexus between compute power, AI/ML/CV, and increased availability of imagery.

Today there are few ways in which to automate object-based image analysis or land-cover classification (Maggiori et al., 2017). Classification is more likely done on an image by image basis. This allows for individual control when classifying an image, but it fails to attain a bigger picture over space and time. Due to the decreased cost of imagery, there is more imagery available to generate analysis over time. The derived large-scale data and analysis can be incorporated into financial, defense, or humanitarian assistance models that deal with spatial data. (Hope, 2016). Instead of working manually

with many images, automation allows for the analysis of thousands of images in the same amount of time. The data can then be incorporated into a broader model or workflow (Buchen, 2015).

To show the value of AI/ML/CV, cloud computing, and increased availability of imagery, different humanitarian and defense use cases are presented below. Section 2.1 examines the Battle of Marawi. The battle took place in the Philippines between May and October 2017. The automated detection of objects at scale over the area of interest (AOI), land cover classification, and insight into geolocation data will be used to show the importance of such data to a military commander or humanitarian organization. Section 2.2 evaluates the automation of multi-class aircraft at scale for peacetime indications and warnings (I&W). The importance for an analyst revolves around the movements of near-peer adversaries, and the time savings associated with AI/ML/CV when compared to a human analyst. The last use case examines the Camp Fire which took place in Fall 2018. A land cover examination shows how large-scale data is used to help guide disaster relief efforts both in near real-time and for post-disaster recovery.

This demonstration of novel workflows seeks to show how the above use cases can benefit from AI/ML/CV. These workflows can be reproduced for future imagery and other forms of geospatial analysis in the geographical sciences, both broadly and for specific humanitarian and defense needs. The object detection algorithms were trained on Digital Globe's Worldview. The land cover algorithm was built and trained on Planet Dove constellation imagery. Use of imagery and geolocation/telemetry data was made possible by Orbital Insight, a geospatial analytics company based in Palo Alto, CA ("Orbital Insight" n.d.).

Chapter 1: Technology Advances

1.1 Increase in Data

New commercial satellite imagery providers, and the cost reduction of building and launching satellites has increased the amount of imagery available (Marr, n.d.-b). This is partially due to venture capital money supporting new companies involved in creating and launching satellites. From 2015 to 2017, it is estimated that a combined $4.2 billion in venture capital investment went to space ventures("The Most Active Space Tech Investors," 2017). In addition, from 2011 to 2016, the number of satellites orbiting Earth increased by about 40 percent (Burningham, 2016). These new datasets, that were once sparse and only limited to government customers, have now been opened for commercial uses.

The growing interest of the data has continued the cycle of building satellites, launching, and then acquiring remotely sensed data for various uses. The private sector has enabled this cycle. SpaceX, a private rocket company, launches its Falcon Heavy Rocket for approximately $90 million per launch ("SpaceX Readies First Falcon Heavy Launch for Paying Customer ," 2019). This contrasts with NASA's Space Launch System (SLS), which costs approximately $1 billion per launch (Powell, 2018). The SpaceX rocket is partially reusable, driving down future launch costs. Therefore, the cost of launching future rockets and imagery satellites decreases. This also leads to an unprecedented amount of data creation in general. In 2018, it was estimated that 90% of the world's data had been created in just the previous two years (Marr, 2018). Although not directly relating to remote sensing data, it provides a metric for understanding the future supply and demand for all forms of data. As seen in figure 1 below, there has been a sudden increase in civilian operated satellites within the last decade (Belward & Skøien, 2015).

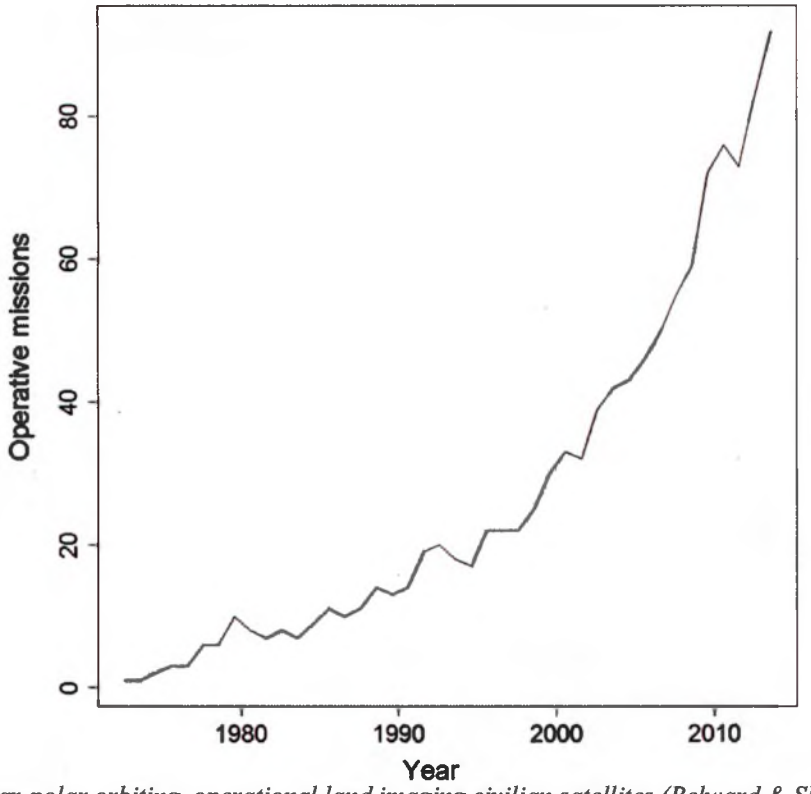

Figure 1. Near-polar orbiting, operational land imaging civilian satellites (Belward & Skøien, 2015).

Since 2010, an increased rate of commercial satellites have been launched. Many of these satellites include multispectral and synthetic aperture radar (SAR) capabilities. Some also are referred to as "constellations," or "small sats." These are groups of satellites that are relatively small, and collectively allow for a higher revisit rate of the Earth's surface ("eoPortal - Earth Observation Directory & News," 2019).

Provider	Satellite Name	Launch Date	Spatial Resolution
Airbus/Astrium	Pleiades 1B	12/2/12	0.5 m (panchromatic) 2 m (multispectral) 0.5 m (pansharp)
Airbus/Astrium	Pleiades 1A	12/16/11	0.5 m (panchromatic) 2 m (multispectral) 0.5 m (pansharp)
Airbus/Astrium	Spot 6	9/9/12	1.5 m (panchromatic) 6 m (multispectral) 6 m (merge)
Airbus/Astrium	Spot 7	6/13/14	1.5 m (panchromatic) 6 m (multispectral) 6 m (merge)
DigitalGlobe	WorldView 1	9/18/07	0.5m (GSD panchromatic)
DigitalGlobe	WorldView 2	10/8/09	0.46 m (GSD panchromatic) 1.85 m (GSD multispectral)
DigitalGlobe	WorldView 3	8/13/14	0.31 m (GSD panchromatic) 1.24 m (GSD multispectral) 3.70 m (GSD SWIR) 30.00 m (GSD CAVIS)
DigitalGlobe	WorldView 4	11/11/16	Panchromatic Nadir: 0.31 m Multispectral Nadir: 1.24 m 65° (earth limb): 14.00 m
Planet	Planet Doves	2014	3 - 5 m
Planet	Sky-Sat	11/21/13	0.9 m (panchromatic) 2 m (multispectral)
Blacksky	Pathfinder-1	9/29/17	1 m
Iceye	ICEYE-X1	1/12/18	SAR
Capella Space	Denali	12/3/18	SAR

Table 1. Sampling of current imagery satellites.

Table 1 is not a comprehensive list of commercial satellites launched in the last ten years ("Satellite Missions Directory - Earth Observation Missions - eoPortal," 2019). However, it shows the increasing number of satellites launched over the last seven years and an indication for more launches in the next decade ("Satellite Launches to Increase Threefold Over the Next Decade - Via Satellite -," 2017). Planet, Iceye, Blacksky, and Capella also are planning on launching more constellations (Spaceflight, n.d.). These small and relatively cheap satellites, combined with commercial launch vehicles, are contributing to the increase of data and reduced cost of imagery. For instance, on 1 April 2018 the India Space Agency delivered 24 U.S.-made small satellites("India space launch: One rocket, 29 satellites, three orbits," 2019). Satellites from Lithuania, Spain,

and Switzerland also were delivered on the same launch, at varying orbits. Government databases also are useful for humanitarian and defense uses. As of 2016, the Earth Science Data and Information System (ESDIS) held 7.5 petabytes of data, all of which was in-domain remote sensing data (Chi et al., 2016). The increase of data has provided the opportunity for exploratory analysis to be conducted. Exploratory analysis on large quantities of satellite imagery may not have been the reason for the increased number of launches carrying imagery satellites. However, understanding what insights and signals can be derived from this type of "big data" is worthy of evaluating. AI/ML/CV and cloud computing is one way to analyze these datasets.

1.2 Cloud Computing

Private satellite and launch vehicle companies have driven down the cost of satellite imagery (Peng, n.d.). Therefore, governments are not the only organizations to have access to this data. As more remotely sensed data is created, the need for storage and analysis also increases. Like the creation of data through multimedia, big data also relates to geospatial data. This can be in the form of imagery, geolocation data from cell phones, and automated identification system (AIS) data for ships. Regardless of what constitutes big data, there are similarities among them. Three characteristics that help explain big data in general, and satellite imagery and other geospatial datasets in particular are: 1) data is numerous, 2) data cannot be categorized into regular relational database, and 3) data are generated, captured, and processed rapidly (Hashem et al., 2015).

A major concern is that data are being generated quicker than they can be processed. This is not a necessarily new phenomenon, but the current scale and diversity of data sources at which it takes place creates new challenges. These challenges span many industries, but it is especially relevant to the geospatial and remote sensing communities due to the novel nature of its applications for humanitarian and defense

issues. Big data can also be explained as "the amount of data just beyond technology's compatibility to store, manage, and process efficiently" (Manyika et al., 2011). Additionally, big data can be understood through volume, variety, velocity, and value (Gantz et al., 2011). This definition also applies to remotely sensed and spatial data. Volume refers to the amount of imagery, in bytes, that is continuously being created. Variety refers to the type of data collected. For the remotely sensed data and spatial data fields, this refers to the different sensors that collect data, including but not limited to private and government sources of multi-spectral, hyperspectral, radar, geolocation/cell phone telemetry, and AIS data. Velocity is the speed of transfer. With advancements of downlink stations, imagery and cloud-based companies now need to create this technology to keep up with the amount of data created. For example, Amazon Web Services (AWS) have introduced ground stations to quickly downlink imagery from satellites ("AWS Ground Station – Ingest and Process Data from Orbiting Satellites," 2018). Lastly, the above needs to have value derived from them in order to prove useful. This means being able to find insights and unobvious value from large datasets. For satellite imagery, it is possible to use cloud computing to analyze large amounts of imagery and attain an understanding of what is happening over space and/or time. Before cloud computing (and eventually AI/ML/CV), this was done manually one image at a time. This degraded an analyst's ability to incorporate a contextual narrative into their analysis.

Cloud services and big data methods therefore work together by utilizing distributed storage technology in the cloud, rather than a local server or device. Big data are evaluated by cloud based applications, often as a service model, reducing the cost for a user who otherwise would need an architecture of their own (Hashem et al., 2015). Cloud computing allows for the processing of data over the internet by providers using "instances." (Jermain et al., 2016). An Elastic Compute Cloud (EC2) instance from Amazon is one such example. An EC2 instance is a virtual server within the EC2 for

running applications on an AWS infrastructure ("Amazon EC2 Instance Types - Amazon Web Services," n.d.). For image processing, it is possible to not only use central processing units (CPU), but also graphics processing units (GPU) in a cloud instance. Previously, GPU-based computing required special hardware from such companies as NVIDIA (Buonaiuto et al., 2017). With cloud computing, GPUs, and parallel computing, new methods of analyzing big data exist. Microsoft, Amazon, Google, and other large technology companies offer their service as a relatively cheap way of using computing power (Sharma et al., 2012).

Large vector and raster-based spatial datasets have also benefitted from the use of cloud computing and GPUs. A simple map projection transformation in a geographical information system (GIS) may not succeed if the dataset is too large. The efficient conversion of large vector datasets from one projection to another is possible through cloud computing and GPU technology (Tang & Feng, 2017). Applying GPU technology and cloud computing to imagery similarly helps computer vision algorithms complete a large-scale task.

1.3 Artificial Intelligence

Artificial intelligence is a broad term. For the purpose of this paper it is used as it relates to computer vision algorithms' application to remotely sensed imagery. Object detection and land cover classification algorithms are applied to imagery at scale in the use cases in Chapter 2.

Machine learning is one part of artificial intelligence. The design of algorithms is based on data that a machine can read, and can incorporate different types of algorithms such as neural networks, decision trees, or random forests (Lary et al., 2016). Machine learning can be used for classification of objects or pixels and be based on a few or thousands of variables. With the increased availability of satellite imagery, a more robust training dataset is possible to train and test algorithms. For instance, an algorithm for

detecting cars ideally uses thousands of training images during development. Training data is needed, and cars would be "tagged" or "labeled" based upon the specific circumstances and parameters of the future algorithm. The algorithm is then "trained" on these images and told what a car is and what a car is not, considering all pixel values for an image. This is repeated, and the algorithm is "tested" on a subset of images it has never seen before. Algorithm performance can then be scored. Typically, algorithms are scored based on a "precision and recall curve." This curve is one way of evaluating algorithm performance. The curve is based on errors of omission and commission, but for computer vision as opposed to remote sensing purposes, precision and recall is used (Boschetti et al., 2004). Figure 2 is a confusion matrix example in which the concepts of precision and recall are taken.

		Predicted	
		Negative	**Positive**
Actual	**Negative**	True Negative	False Positive
	Positive	False Negative	True Positive

Figure 2. Confusion matrix example.

Precision (Equation 1) relates to the metric measuring "out of everything the algorithm detected, what percentage was actually what it wanted to detect." Or, "ability of a classification model to return only relevant instances." (Koehrsen, 2018).

$$Precision = \frac{True\ positive}{True\ Positive + False\ Positive} \qquad (1)$$

Recall (Equation 2) means "out of all available objects that should have been detected, what percentage were actually detected." Or, "ability of a classification model to identify all relevant instances (Koehrsen, 2018).

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \qquad (2)$$

These two numbers create a precision and recall curve. Algorithms can also be built to increase the weight of one of these metrics. So, if an algorithm has very high precision, it may have a lower recall. Conversely, if an algorithm recall is high, it may also have a higher number of false positives. Therefore, there are tradeoffs when trying to optimize for one performance metric. Optimizing the metrics that are most effective for a specific problem will vary. Lastly, the F1 score (Equation 3) is a single metric that averages precision and recall metrics for a model.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \qquad (3)$$

Specifically, this paper will showcase computer vision algorithms being used at scale, analyzing thousands of images in order to understand macro level trends that help guide more pointed investigation. These algorithms are part of a web-based geospatial analytics platform created by Orbital Insight, Inc.

1.4 Algorithm Details

The car detection algorithm was built on Worldview imagery from Digital Globe, with parking lots in the United States collected as training data. Imagery was then tiled and uploaded to a web-based data labeling software. Technicians then "tagged" thousands of cars in the imagery to create a training set. Figure 3 shows initial algorithm performance on a test set of imagery showing cars in parking lots. Initial car detector

performance was evaluated to be above 80% precision and recall. Future improvement includes training the car detector algorithm on imagery other than parking lots or paved surfaces. Dirt, grass, and other ground surfaces should be implemented into future training sets to improve overall algorithm performance on non-pavement surfaces.

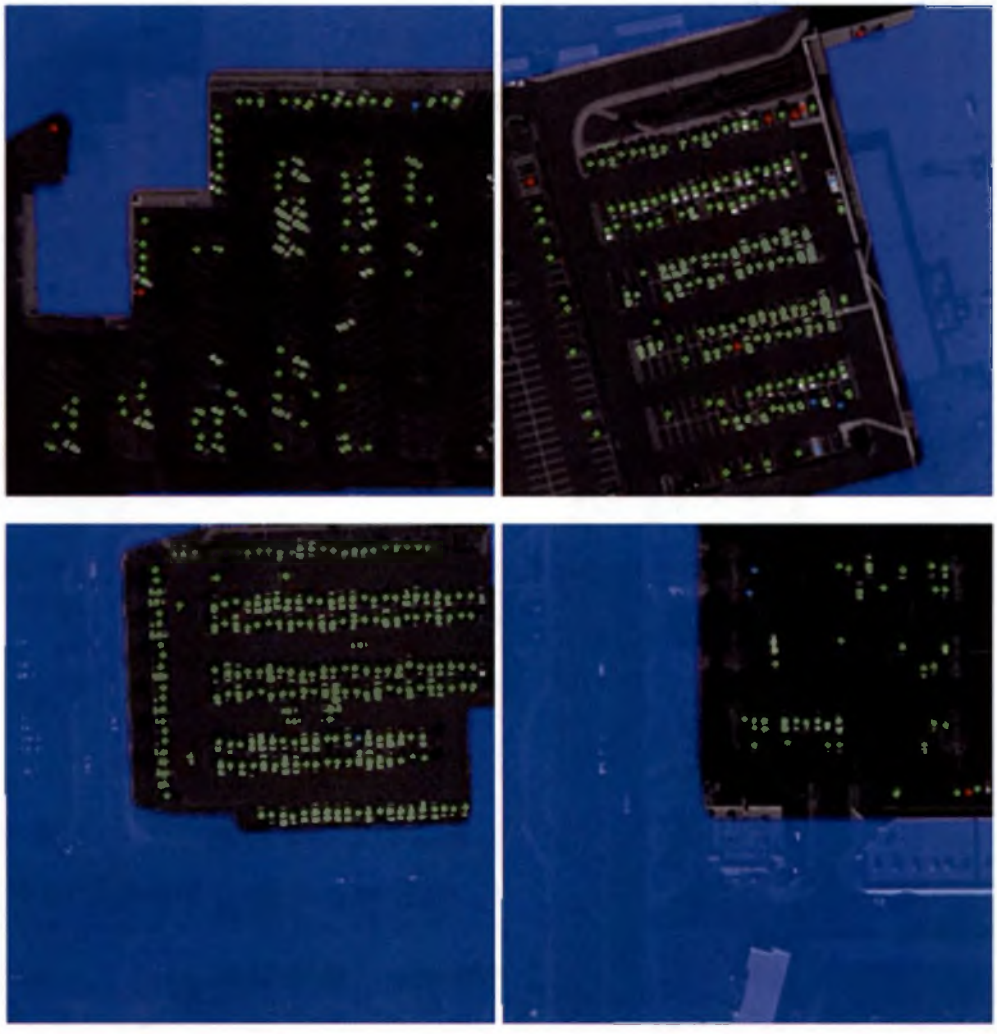

Figure 3. Initial car detector algorithm performance. Green dots are true positives, blue dots are missed detections, and red dots are false positives. Courtesy of Orbital Insight.

The land cover classification algorithm used mostly imagery from urban areas around the globe. Due to the imagery that is available in Planet's catalogue, imagery sampling skew towards urban areas. Like the car algorithm, the land cover classification algorithm was built on training data created by human image analysts. Thousands of images are tiled and uploaded into a proprietary marking tool. Classes can then be "painted" or "tagged" based on the customized marking instructions provided. The land cover class primary focused on the building and road classes during the marking campaigns. This ultimately lead to higher performance for the building and road classes versus the other classes. Future improvement would include gathering more rural and non-urban training datasets. Precision and recall scores vary by class, but as seen in figure 4 initial land cover algorithm output visualizations suggest relatively valid results.

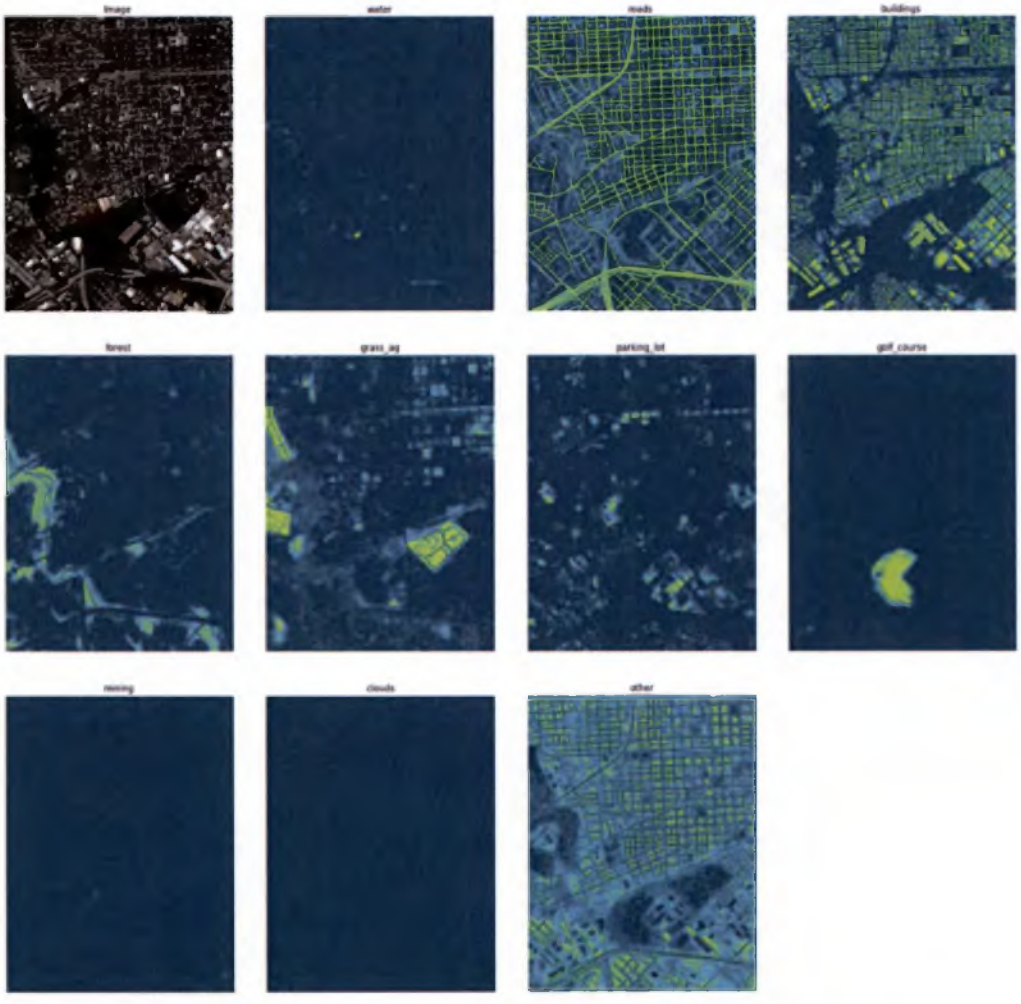

Figure 4. Initial output visualizations from land cover classification algorithm broken down by class. Courtesy of Orbital Insight.

Like the car detection and land cover classification algorithms, the multi-class aircraft detector went through a similar process during development. Imagery was gathered from airfields around the world, tiled, and uploaded into a web-based marking software tool. Aircraft were then manually marked using a customized marking tool within the marking software. In phase one, specific points on the aircraft were marked. During phase two, the marked aircraft were then categorized as fighter, bomber, commercial/passenger, and other. Machine learning was used to train an algorithm, which

was then tested on a different set of test imagery, not yet seen by the algorithm. Identifying true positives, false positives, false negatives, and false positives helped quantify the initial performance of the algorithm. This was also done for the car detection and land cover classification algorithms. Performance varied by class depending on two main factors. The first was how many aircraft of a specific class was marked. The more training data, the more likely that class will have a higher precision and recall. Second, how distinct an aircraft looks helps create a standardized type of object the algorithm can look for. For instance, a fighter plane looks very different than a commercial plane. Therefore, the algorithm has an easier time distinguishing the two types during the machine learning process. This contrasts with the difficulty the algorithm may have when trying to figure out the difference between a large cargo plane and a bomber plane. For the classes used in the use cases, precision and recall were all above 70%. Figure 5 shows a sampling of true positives during initial algorithm performance on a test dataset.

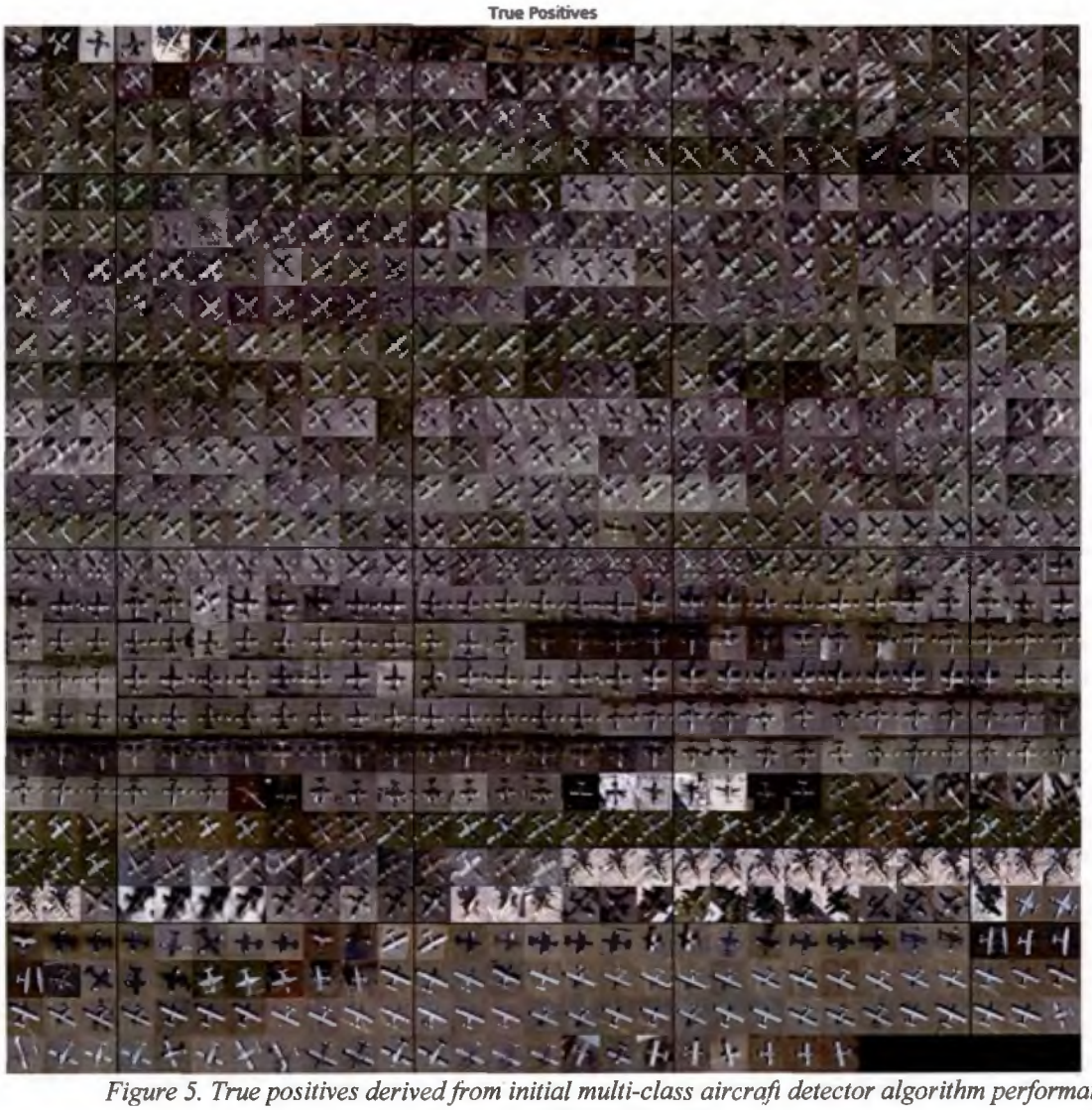

Figure 5. True positives derived from initial multi-class aircraft detector algorithm performance. Courtesy of Orbital Insight.

Chapter 2: Use Cases

The use cases to follow and the associated technology are intended to show the benefits of AI/ML/CV when it comes to large-scale analysis. Example personas that would be able to use these workflows for their benefit include humanitarian organization personnel, military commanders, or intelligence analysts. With limited time and

resources, they are most likely to have to focus on finding and reporting on visible phenomena immediately relevant to the situation: military vehicles/equipment, barricades, closed neighborhoods, physical damages, infrastructure status, etc., leaving little to no availability for second-order impacts or locations, or longer-term predictive considerations. The introduction of CV-derived object detections either at selectively indicative locations such as destroyed buildings or at scale across the entirety of the city, can provide an auto-derived set of data with which to enrich the otherwise (justifiably) myopically-focused conflict reporting. Consequently, this provides a far greater depth of analysis and volume of quantitative data to the geospatial/imagery analyst's reporting, in turn increasing its utility to external analyst counterparts or national decision-makers.

For the Battle of Marawi, these personas would either already be on the ground or preparing to intervene or monitor the battle at hand. Car counts, land cover changes, and geolocation information all help to provide personas with timely information that helps guide their mission decisions.

By monitoring airfields globally, a geospatial/imagery analyst responsible for covering the region or subject matter has a responsibility to provide imagery-based reporting on aircraft counts at various airfields. Logging these counts creates a historical trend. From these trends, "Indications and Warnings" can be derived and thresholds for anomalous activity decided.

Lastly, large-scale and timely land cover classification will be briefly examined in how it related to the Camp Fire in Northern California during the fall of 2018. This analysis has the potential to help wild land firefighters and other humanitarian assistance /disaster response personnel identify the situation on the ground with remote sensing and CV algorithms.

2.1 Battle of Marawi

2.1.1 Introduction

The Battle of Marawi was a five-month long conflict that started 23 May 2017 and is considered the longest urban battle in modern Philippine history (Singh, 2018). The main belligerents involved included the Armed Forces of the Philippines (AFP) and militants affiliated with The Islamic State (ISIS), including the Maute and Abu Sayyaf Salafi jihadist groups (Franco, 2017). Essentially these terrorist organizations took over part of the city, and the AFP's mission was to clear and re-take the city. By doing so, many human and physical geographical changes took place. Patterns of life changed, civilians left the city, and many of the cities buildings were damaged or completely destroyed (Tagoranao & Gamon, 2017) .

This use case shows tools and workflows that an analyst could deploy to produce information on a conflict area. First, car counts in the most heavily affected area of the city will be shown. Next, data from a land cover algorithm will be used to create visuals that show changes in building and road land cover classes. Lastly, sample geolocation data for the Marawi area will be shown against that of a nearby city where fighting did not take place. The computer vision algorithms used were built by Orbital Insight.

2.1.2 Marawi Car Counts

Digital Globe high-resolution imagery was the sole provider utilized for car detection. This was determined through the high performance and relevance of the DG Car Counter to this specific analysis. Once imagery is ingested into the pipeline, and tiling and pre-processing are completed, algorithms can be processed relatively quickly. The AOI examined is the southeastern part of Marawi, which is separated from the "safe zone" to its northwest by the Agus River and bordered to its south/southwest by Lake Lanao. This area was most affected by the fighting (Gunaratna, 2017).

Approximately 148 scenes were analyzed and 35,000 cars were counted. Setting up the project parameters took approximately 15 minutes. It took about 24 hours of compute time to order, ingest, and analyze all imagery. Without automated ingestion pipelines and car detector algorithms, the entire process would have taken several days. This would make the nature of the data less useful if it was needed by someone who is doing daily analysis as the battle unfolds. It should be noted that the larger the AOI, the longer it takes for all required imagery to be ingested. Also, note that due to the geography of the region, many scenes were unusable due to cloud cover. Figure 6 shows an aggregated hexbin plot of cars that were detected using car detector algorithm from January 2009 to June 2018. This map illustrates the scale and speed in which the data can be created, as well as an initial look into dispersion characteristics. Digital Globe imagery was used with a 50% cloud cover filter when querying available imagery.

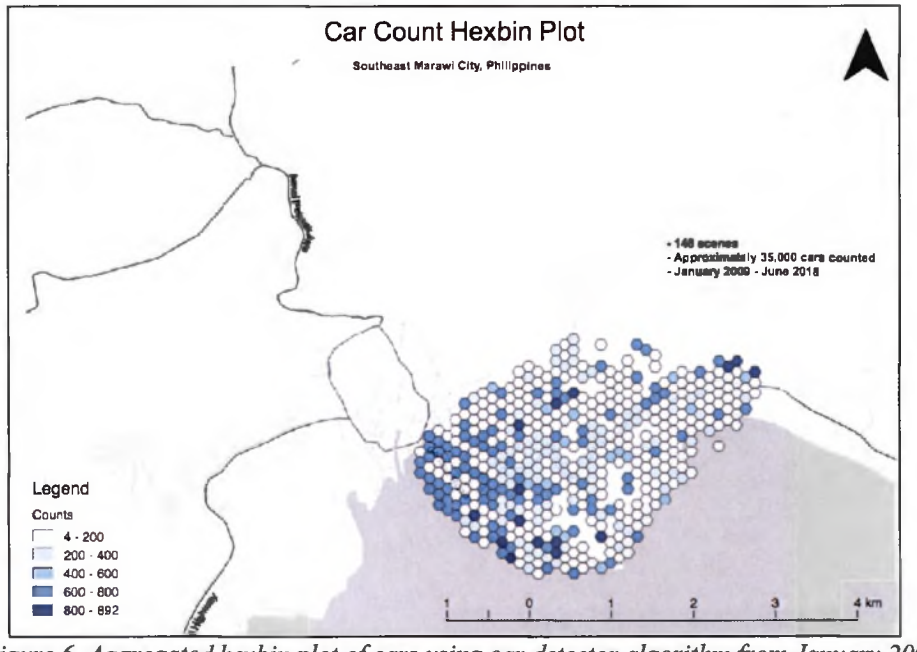

Figure 6. Aggregated hexbin plot of cars using car detector algorithm from January 2009 to June 2018. Made with QGIS v3.0.

Based on the time series of this AOI, we note that raw car counts immediately decline once the battle starts on 23 May 2017. These trends are consistent with the expectation that commercial and passenger traffic through Marawi are severely disrupted due to the battle. The severe drop in automated car counts as soon as the battle begins would warrant a more in-depth analysis. Alerts for car count increases and decreases can then be set up based on the historical counts if needed using specific thresholds. Upon receiving a "decline in car counts" notification, an analyst would be prompted to examine the area more closely. The key results of the analysis include: 1) Notable drops in raw and rolling mean values during the initial start of the battle in late May 2017 and 2) As of July 2018, car counts have not completely recovered to their pre-battle numbers, possibly indicating that the city is still rebuilding, lacking infrastructure, and not suitable for repopulation. However, there is some hint of car numbers starting to increase after the battle possibly indicating cleanup and reconstruction. These key results can be seen when looking at figures 7 and 8. Car counts were derived from the available Digital Globe imagery dating back to the beginning of 2009.

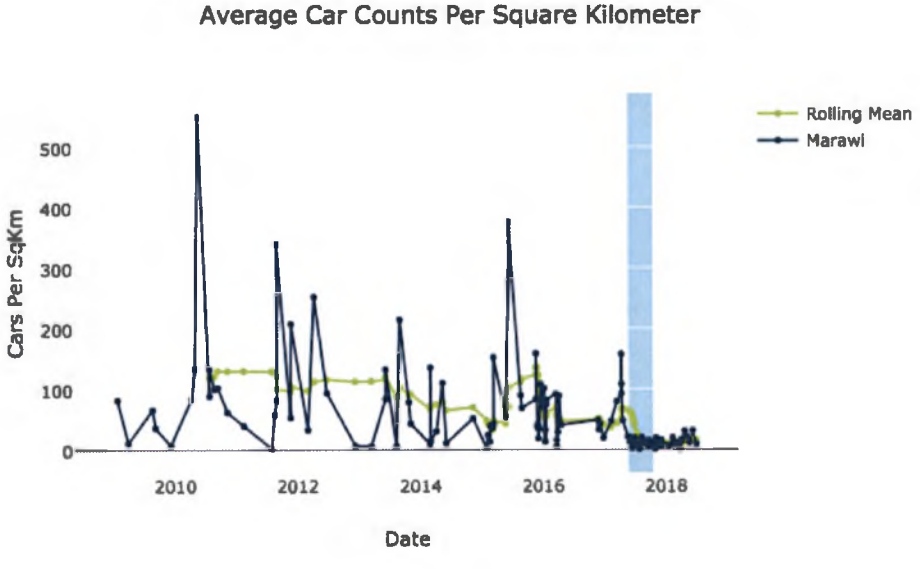

Figure 7. Car counts per square kilometer for the Battle of Marawi AOI. The light blue block indicates the duration of the battle. Made with Plotly.py library.

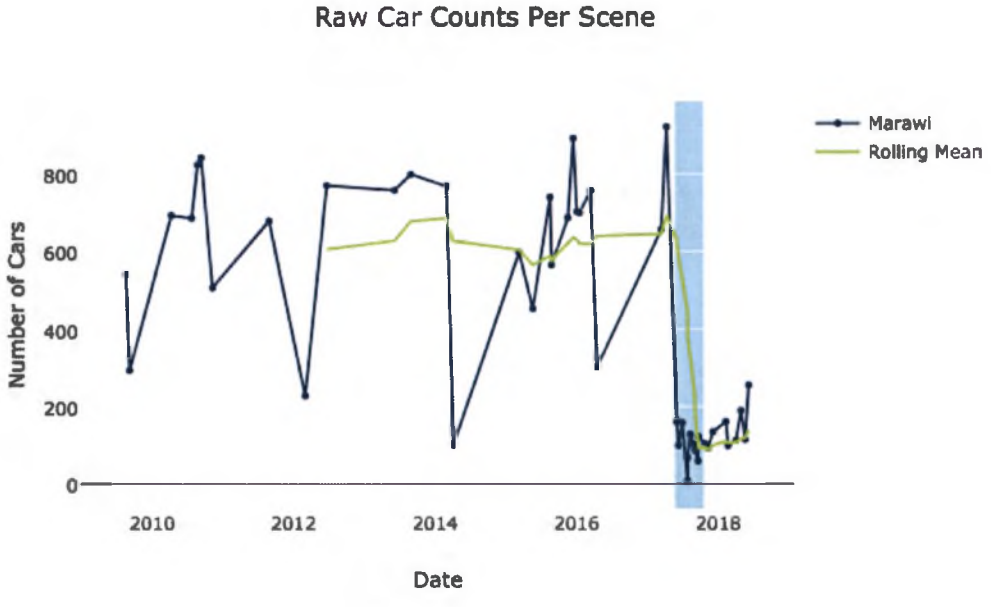

Figure 8. Raw car counts for the Battle of Marawi AOI. The light blue block indicates the duration of the battle. Made with Plotly.py library.

Due to the relatively large size of the AOI, not all scenes provide 100% coverage as seen in figure 9. Conversely, figure 10 shows a scene covering the entire AOI. A threshold for the percentage of AOI unobserved was set to a minimum threshold of 25% coverage. This allows for the customization between using more scenes and car count observations, even if the scene only covers a small part of the AOI, versus less scenes that may have a more complete coverage profile. Therefore, it is likely that the very small counts shown are the result of limited coverage scenes, or scenes that are only covering areas that normally have low car counts (a field versus a parking lot). In this case, the rolling mean of car counts can be used as a better patterns of life indicator. Average car counts per square kilometer was also used as another way of verifying that car counts did

in fact decrease once the battle began. Web-based graphic user interfaces (GUI) are also being created to visualize results detected by the back-end computer vision analysis. Figures 9 - 11 are such examples from an early version of Orbital Insight's platform.

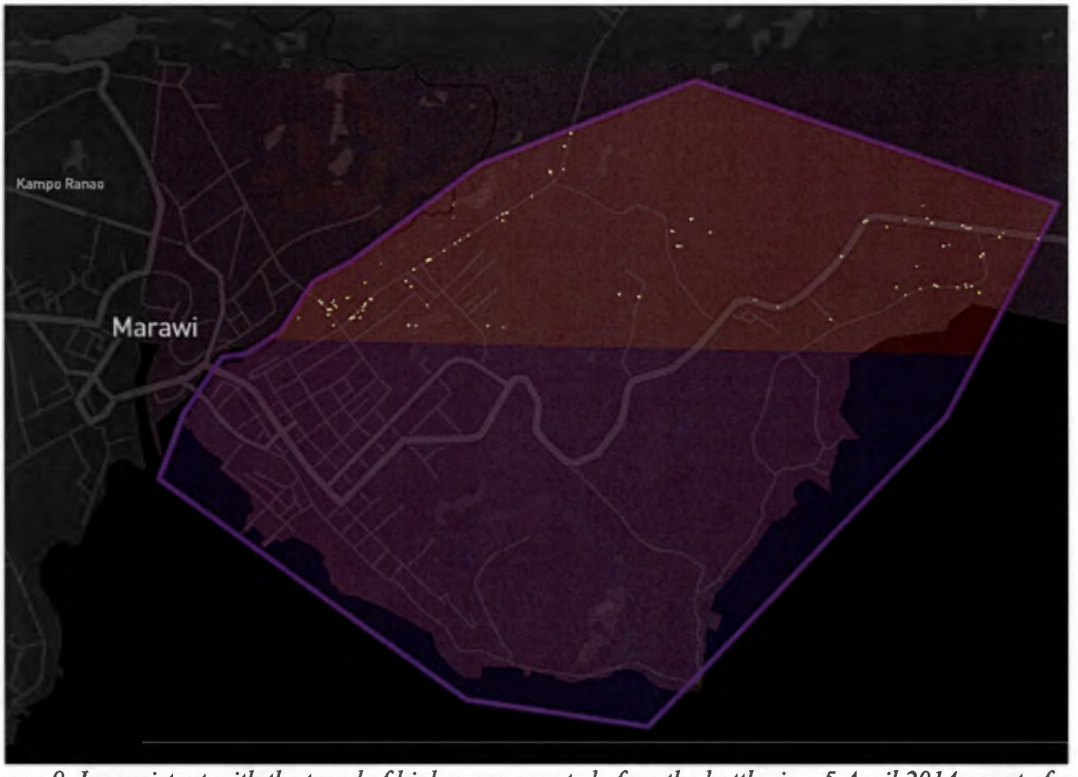

Figure 9. Inconsistent with the trend of higher car counts before the battle, is a 5 April 2014 count of only 100 cars. However, upon closer examination, this scene provided only partial coverage for the AOI. Also, the part of the AOI where there was coverage was in arguably a less densely populated area. These factors can be considered during post project analysis in order to create a more accurate portrait of what is happening on the ground. Map created from Orbital Insight platform using Mapbox basemap.

Figure 10. Cars visualized as dots from an image dated 10 April 2017. 892 cars were counted in this part of Marawi, just over a month before the battle began. This scene had total AOI coverage. Map created from Orbital Insight platform using Mapbox basemap.

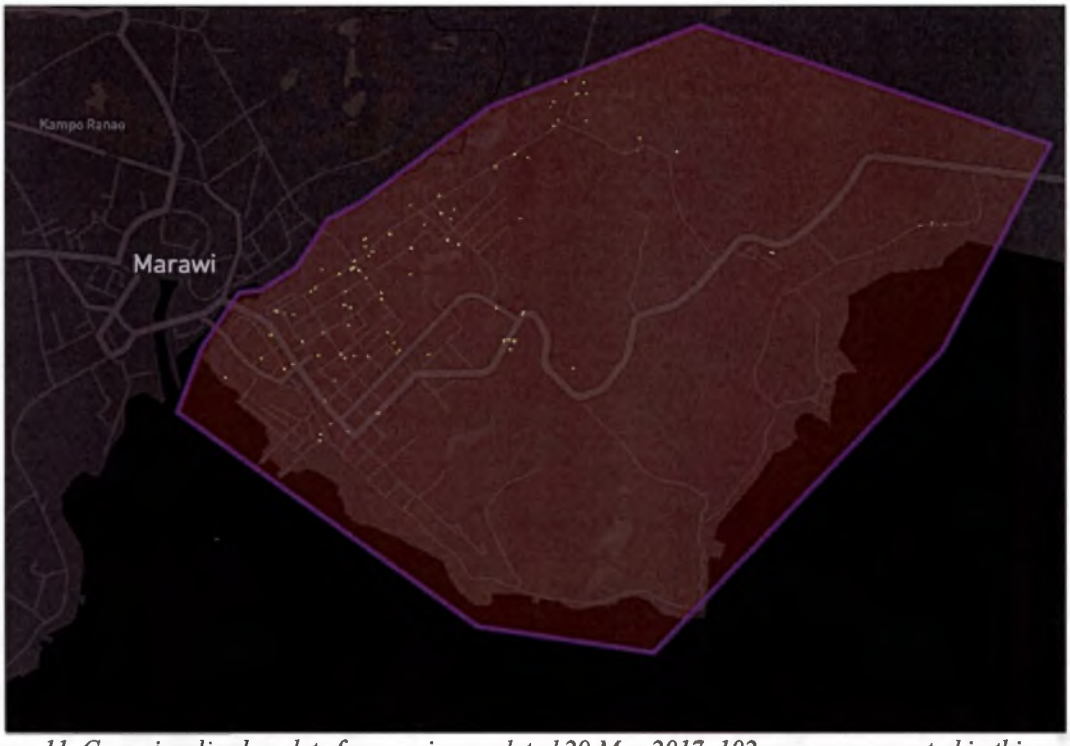

Figure 11. Cars visualized as dots from an image dated 29 May 2017. 102 cars were counted in this part of Marawi a week after the battle began. The scene had total AOI coverage. Map created from Orbital Insight platform using Mapbox basemap.

Because of the potentially randomized nature of single-collect observations, aggregated detection statistics along monthly intervals are examined, as seen in figure 12. In this case, figures are summarized by the maximum and average values of monthly composites. The monthly high and monthly average are helpful in a few ways. When looking at historical imagery in 2009, many monthly highs and monthly means are the same. This would indicate only one scene was taken during that month. Overall, there is a decline in car counts once the battle started. Car counts at scale also can be used to track reconstruction and repopulation efforts post-battle, assuming higher revisit rates continue as compared to pre-battle revisit rates.

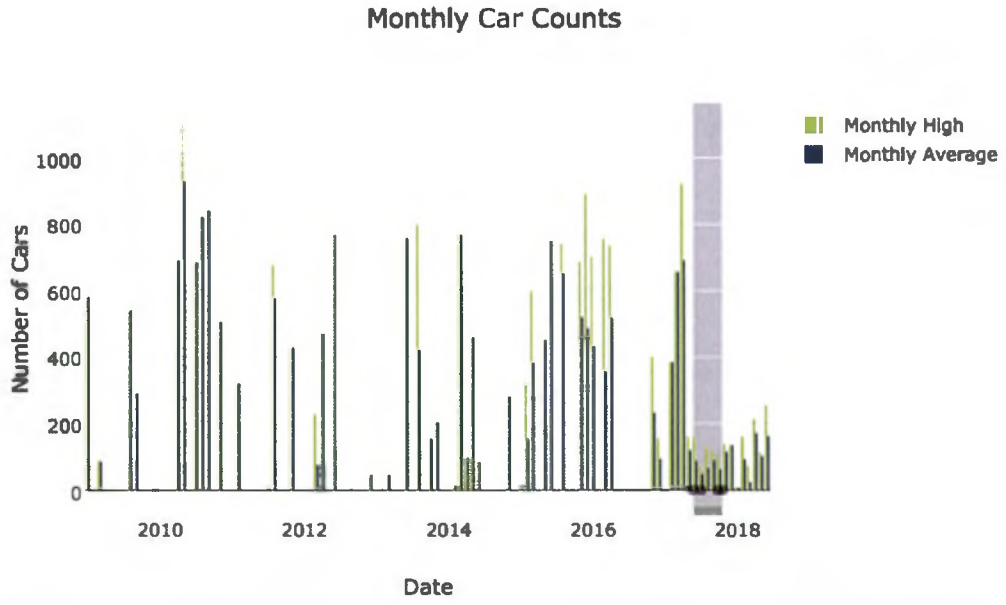

Figure 12. Monthly car counts in the Battle of Marawi AOI. The gray block indicates the duration of the battle. Made with Plotly.py library.

In addition to the overall impact of conflict on traffic within Marawi, spatial patterns are analyzed to derive additional conclusions or open new lines of analytic questioning. An interactive web map displaying commonly accessible background maps underneath CV-derived car points using Orbital Insight's proprietary web-based platform is generated. This allows analysts to display and explore their findings in a readymade graphic format. Prior to the start of the battle, the most densely populated car counts were found in the west/southwest part of the conflict area AOI. Once the battle began, and civilians displaced, the general civilian pattern of life (PoL) changed. Car counts decreased overall within the combat AOI, and their dispersion arguably changed as well. This is especially evident in later stages as well as post battle. If a major city normally has the highest car concentration in the central business district (CDB), and that city was the location of a battle, then it can be assumed that the normal civilian PoL would be

affected. This would mean that there may not be as high of a concentration of cars in the CBD during or after the battle, when fighting and/or reconstruction is taking place. By understanding the counts and dispersion of cars, an analyst can better identify civilian and/or military patterns of life. In this case, the removal of civilian cars in an area of the AOI that normally has the highest concentration of cars could queue the analyst to focus more on the movements of the cars that are still there. Furthermore, if a high concentration of cars returns to what was considered the normal PoL car concentration location, then it could indicate a return of displaced persons. Figures 13 - 17 show the spatial distribution and counts of car detections at varying times leading up to, during, and post battle.

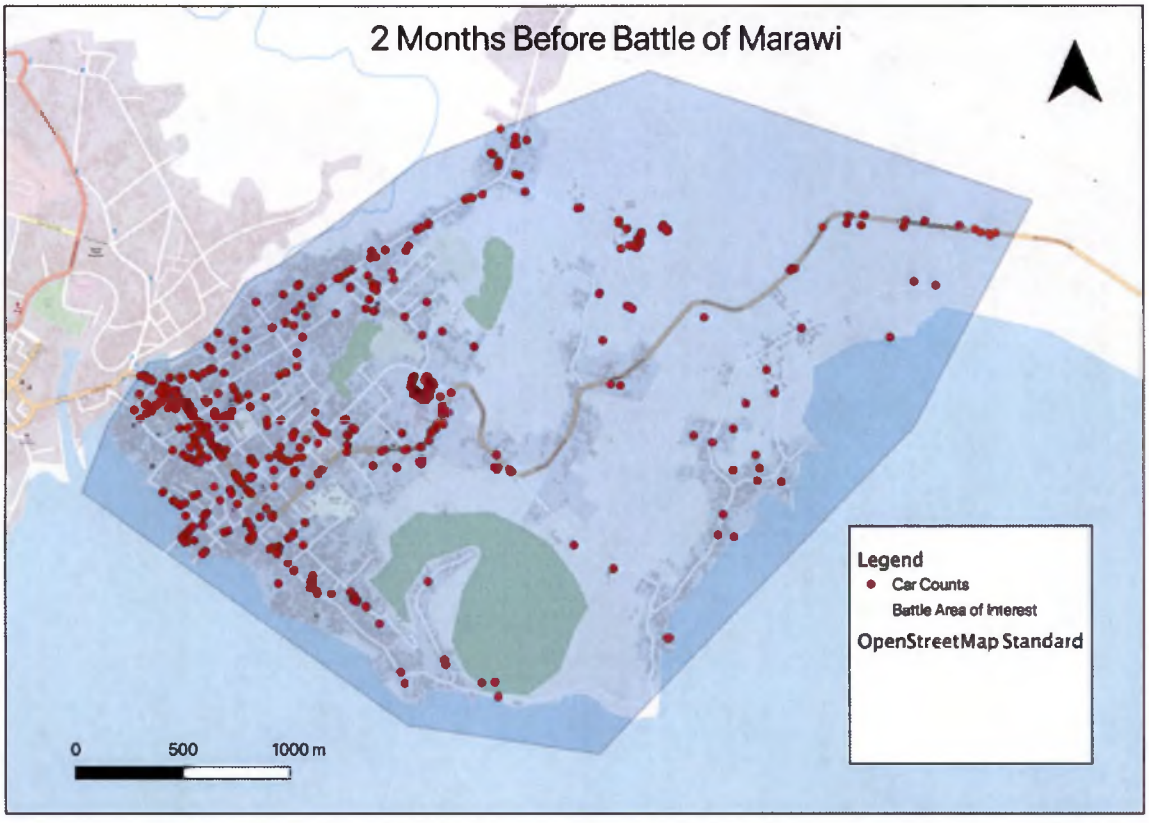

Figure 13. 14 MAR 2017: Before battle; a normal pattern of life with car counts clustered in what is assumed the central business district (CBD). Made with QGIS v3.0.

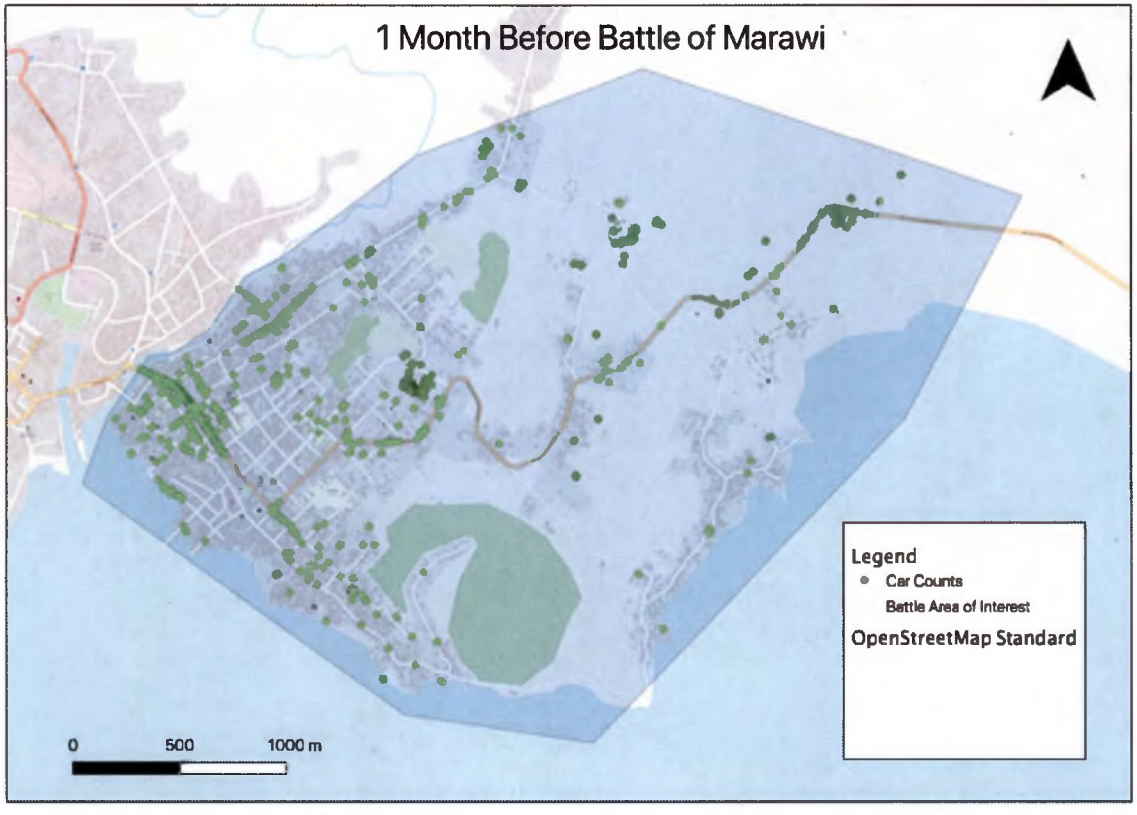

Figure 14. 10 APR 2017: Before battle; another normal pattern of life scene observed, cars still clustered in CBD. Made with QGIS v3.0.

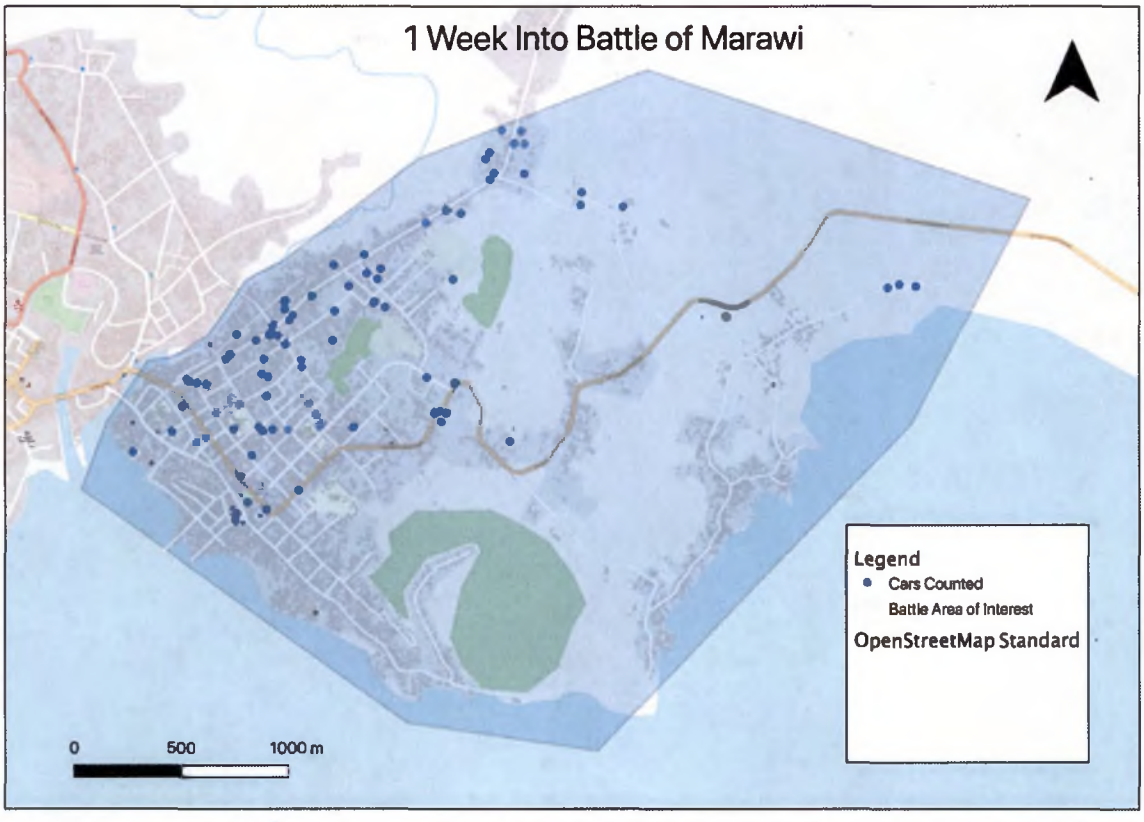

Figure 15. 29 MAY 2017: One week into battle; severe decline in car counts overall, still some clustering in CBD, potential exodus of civilians. Made with QGIS v3.0.

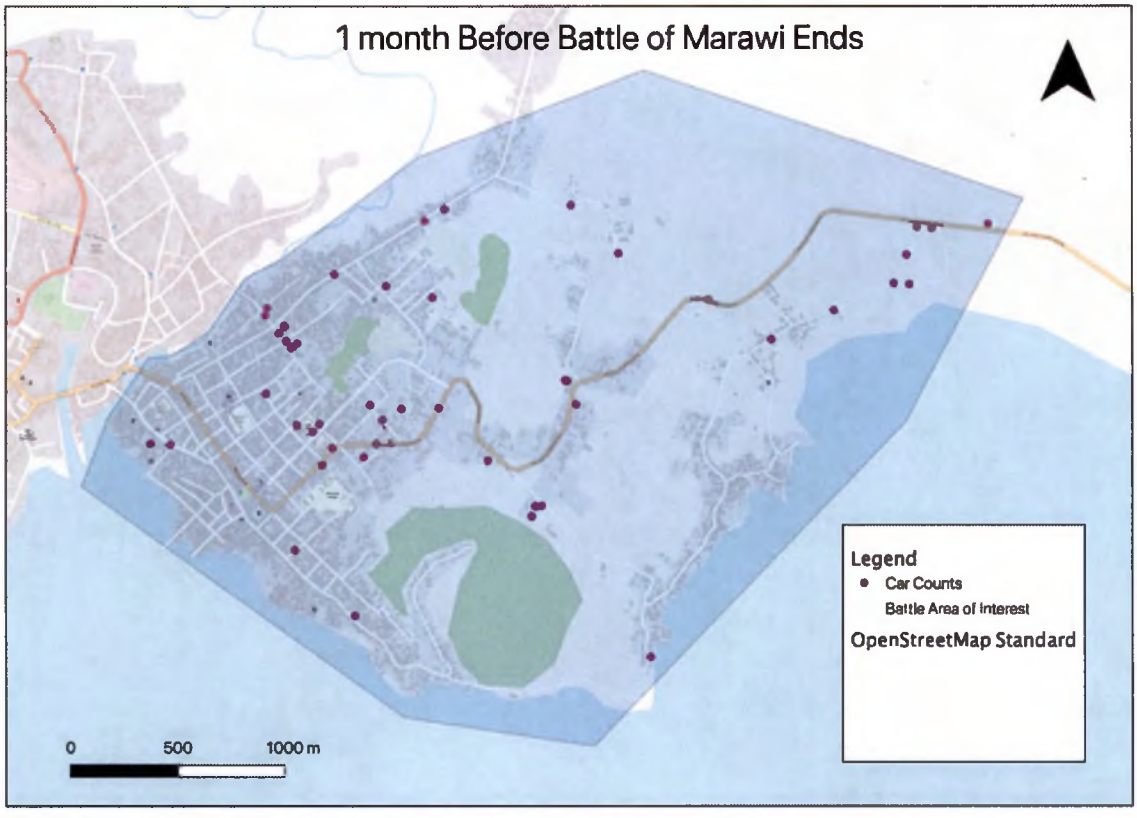

Figure 16. 18 SEP 2017: One month before end of battle; continued decline in car counts, less clustering, potential removal of most civilian vehicles. Made with QGIS v3.0.

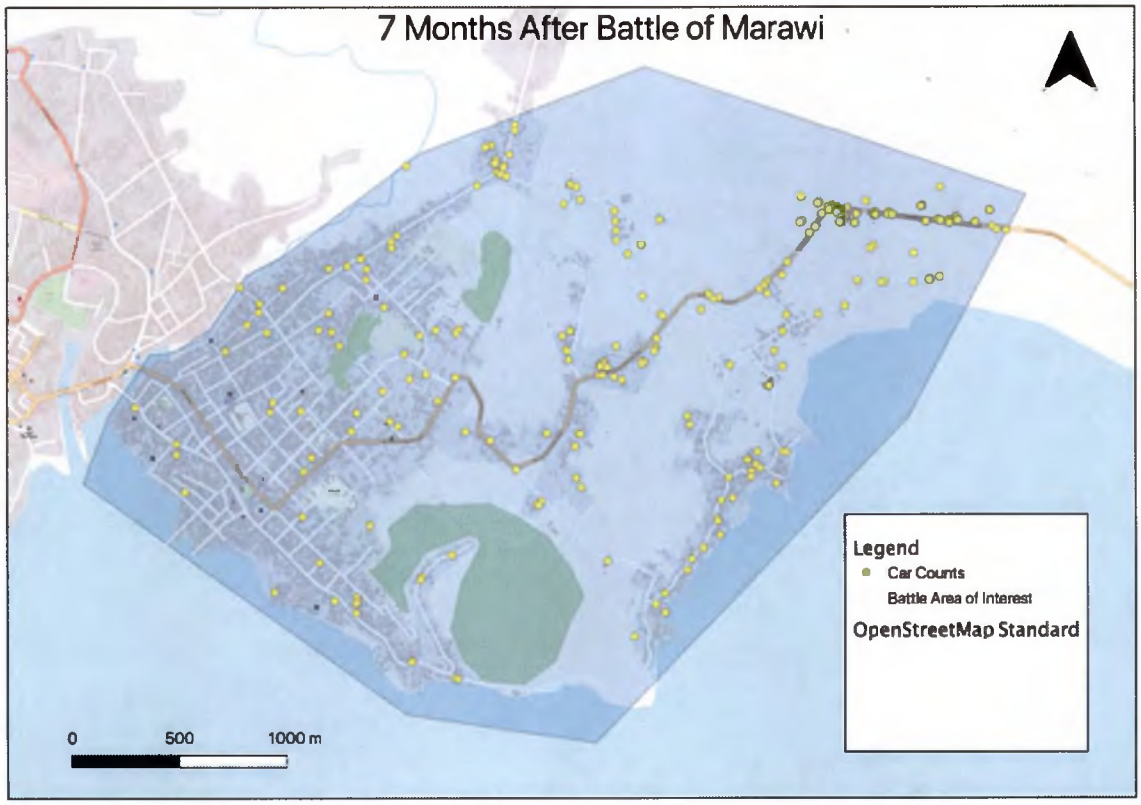

Figure 17. 07 JUN 2018: Seven months after battle; car counts beginning to increase, no CBD clustering, potential reconstruction efforts but no return of civilian PoL. Made with QGIS v3.0.

Sample scenes in the above graphics were chosen based on the best AOI coverage and lowest cloud score. The dates reflect pre, during, and post battle timeframes. With this data, analysts begin to analyze the spatial distribution of cars over time, and make educated guesses concerning what is happening on the ground. Different counts of cars, as well as dispersion and clustering give insight into what is happening on the ground day to day. This data can be implemented into a larger humanitarian or operational plan that allows resources to go where they are most needed (Quinn et al., 2018). For instance, if

less civilian cars are in the city, a humanitarian organization can begin to look outside the city for where displaced peoples are settling, instead of wasting time and resources trying to get into the city to provide aid.

2.1.3 Marawi Land cover Analysis

According to the Office of the United Nations High Commissioner for Refugees (UNHCR), 98% of the population of Marawi City was forcibly displaced (Refugees, n.d.). When the battle was declared over, many of the city's buildings were either destroyed or damaged due to aerial bombing, fire, or other explosive weaponry (Tagoranao & Gamon, 2017). By running a land cover algorithm over available imagery, an analyst can begin to track classification changes. Figures 18 - 20 help military planners and humanitarian relief personnel have a better, high-level understanding of what is happening in their area of operations. These visuals could eventually be incorporated into maps or other tactical mission planning products that are used to brief mission orders. Subsequent mission planning products could then be physically taken on a tactical mission where the status of a building or road is needed and can be easily referenced by the operator. This also proves useful for humanitarian organizations that need updated maps of roads and buildings to guide their efforts of providing aid and reconstruction. Land cover can also help automatically detect displaced-persons camps (Hassan et al., 2018)

The proprietary land cover algorithm available via Orbital Insight analyzed available Planet imagery over a 45 square kilometer AOI between January 2016 and November 2018. By creating a broad timeframe, a user can perform analysis on specific time ranges and classes. Once the classifications are complete over the imagery, the data can be queried to find quantifiable class information. For this use case, buildings and roads classes were used as indicators for what is happening on the ground. To find

before-and-after metrics, a Python script was used to detect changes between a "Baseline" and "Target" time range. The script uses date/time and AOI parameters to query two separate time ranges and the available data, and then determine if changes in class detections for buildings or roads took place. Changes are identified as "destruction" or "construction," indicating a removal or addition of a polygon between the time ranges, respectively. In our case, the land cover algorithm creates class polygons, so the first metric performed simply was a count of "construction" and "destruction" polygons. The time ranges used were one month before the battle began and 13 months after the battle started. This analysis gives us a high-level indication if there were more buildings and roads destroyed or constructed once the battle began. The analysis showed (figure 18) that for these time ranges, 700 buildings were destroyed while only 150 were constructed.

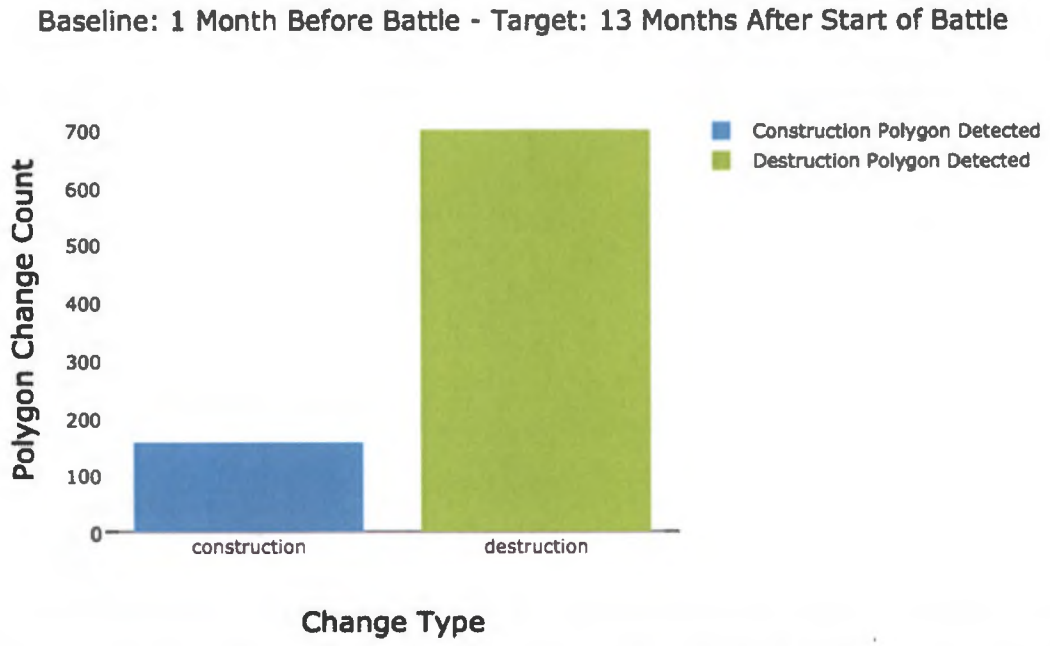

Figure 18. Polygon count changes based on specified baseline and target time ranges for battle AOI. Made with Plotly.py library.

Selecting different baseline and target time ranges helps to understand what is the normal PoL. Baseline and target time ranges over the AOI when the battle was not taking place was used as a starting point in comparison to other time ranges. In figure 19, the output for time ranges when the battle did not take place is shown. All of 2016 was used as the baseline, and January – May 2017 was used as the target time range (right before start of battle). By evaluating different time ranges, change detection is observed for building and road classes. Change detection is not new, and pixel and object based classifications have been used to detect changes (Hussain et al., 2013). However, the scale and speed at which the process takes places allows for the analysis of hundreds of images in the time it would take a human to classify one. The land cover classification algorithm, like other methods, is based on a proprietary convolutional neural network to using spectral values and other parameters (Tewkesbury et al., 2015). Once in the database, a query can be performed to identify class changes between time periods. The output can be visualized as polygons in a GIS. There are relatively few changes in figure 19 as compared with figure 20, which shows changes between one month before and one month after the start of the battle. It would be expected that as the battle progressed, more buildings would be damaged or destroyed. Figure 21 shows how analysis can be portrayed on a rolling basis (month one of battle versus month two). By having a large database of land cover classifications, a user can conduct more specific time range analysis in order to get a better understanding of what is happening on the ground. The user can also reference historical analysis from previous imagery to see how the current analysis is similar or different. This allows the user to better understand PoL, what is normal, and what anomalous. This would be difficult for a user to perform manually at scale on an image by image basis without the use of the land cover algorithm. Instead, the user would most likely analyze a single image at a time manually, without understanding the previous pattern of life of the area.

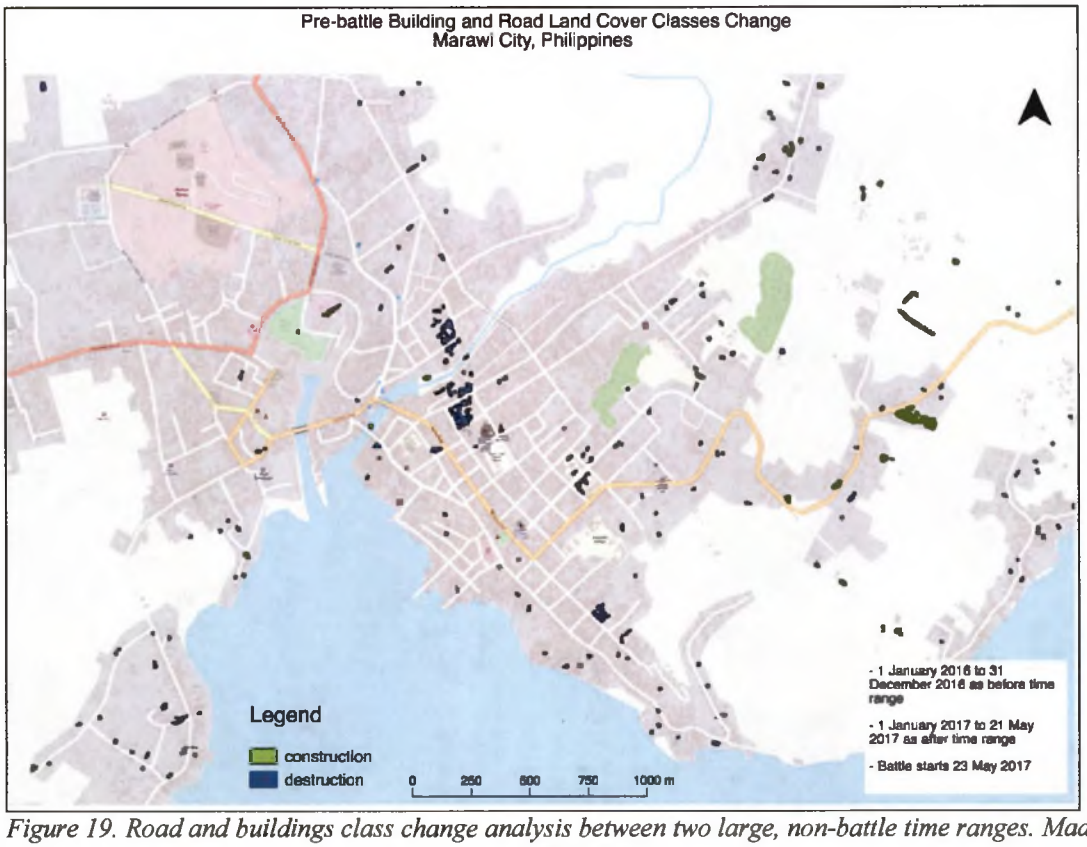

Figure 19. Road and buildings class change analysis between two large, non-battle time ranges. Made with QGIS v3.0.

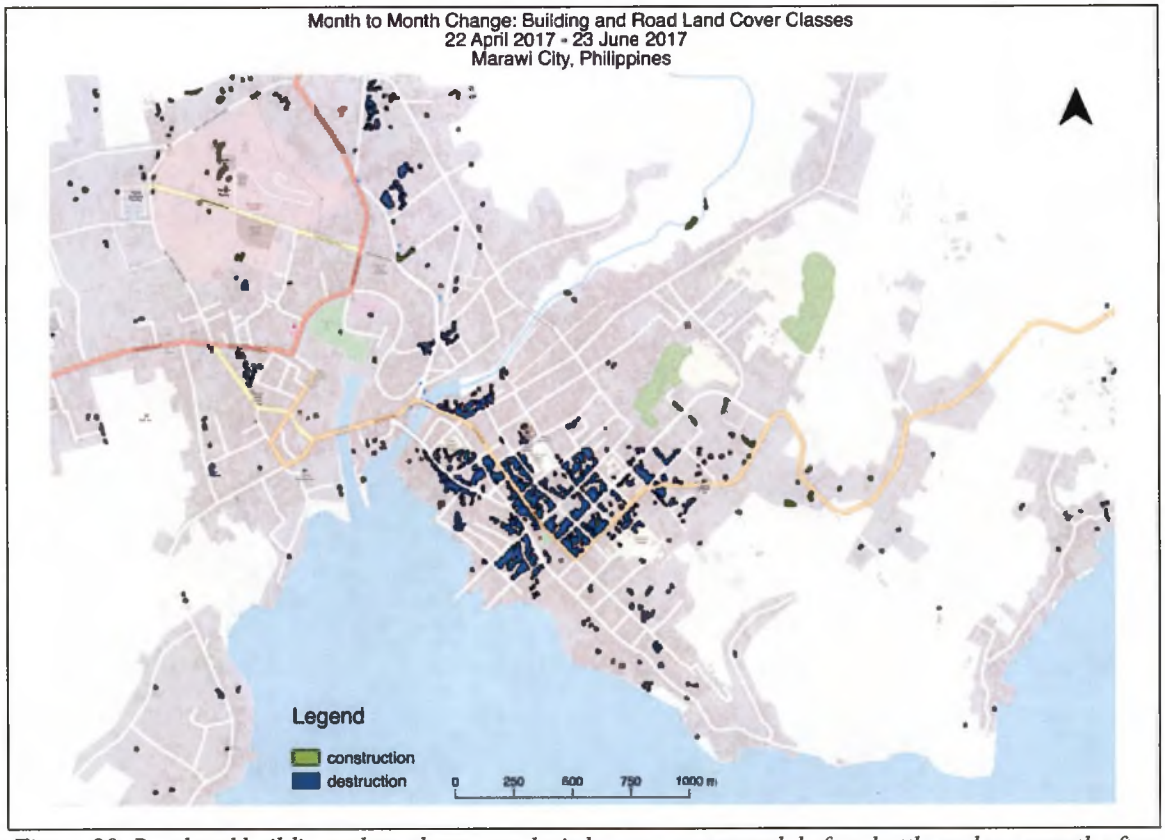

Figure 20. Road and buildings class change analysis between one month before battle and one month after start of battle. Made with QGIS v3.0.

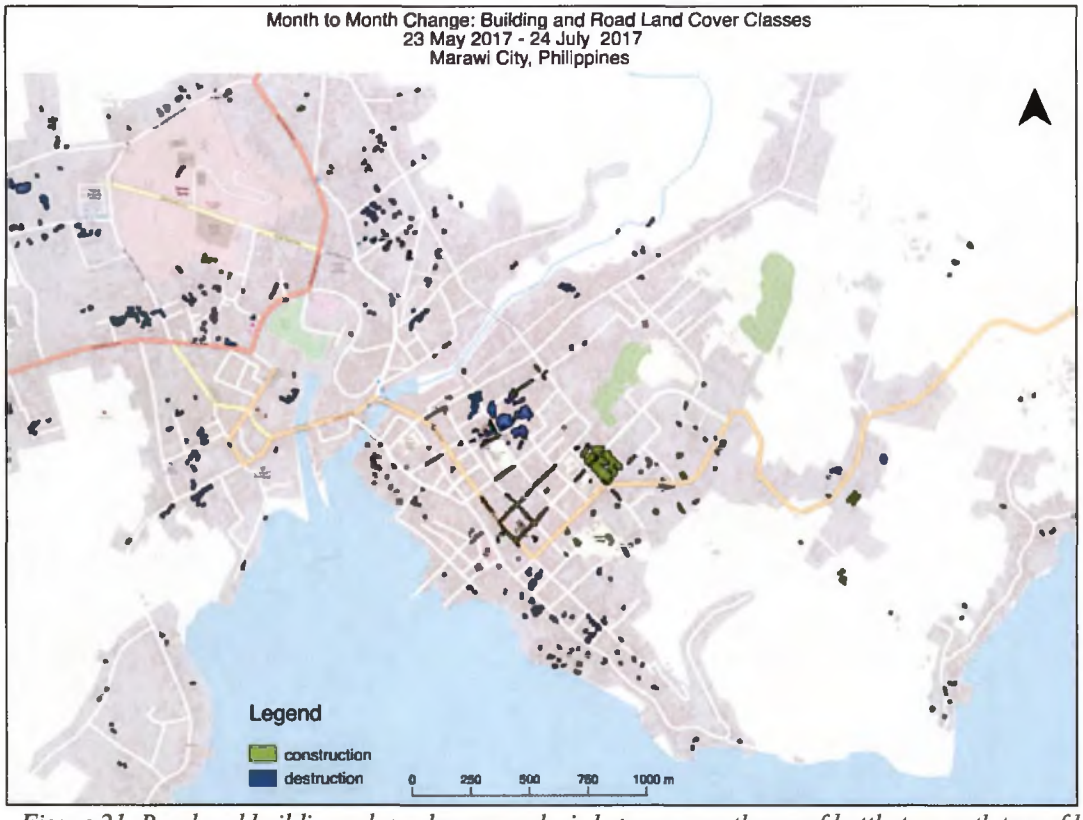

Figure 21. Road and buildings class change analysis between month one of battle to month two of battle. Made with QGIS v3.0.

As fighting continues, more strategic targets are engaged. The group of destruction polygons can be used to indicate happenings on the ground. In this case, an analyst would have guidance to investigate further. When doing so they would find that government forces were closing in on a Maute terrorist stronghold. The stronghold is a mosque and houses the Masjed Mindanao Islamic Centre. A closeup is shown in figure 22.

> The mosque, which houses the Masjed Mindanao Islamic Centre, had been used by militants from the Maute terrorist group as a shelter, stockade, sniper's nest and holding area for their hostages. Satellite images show that the area around it has been reduced to rubble. The army had avoided bombing it out of respect for Muslims." ("Govt troops retake grand

mosque in Marawi, Asia News & Top Stories - The Straits Times," 2017 August, 26)

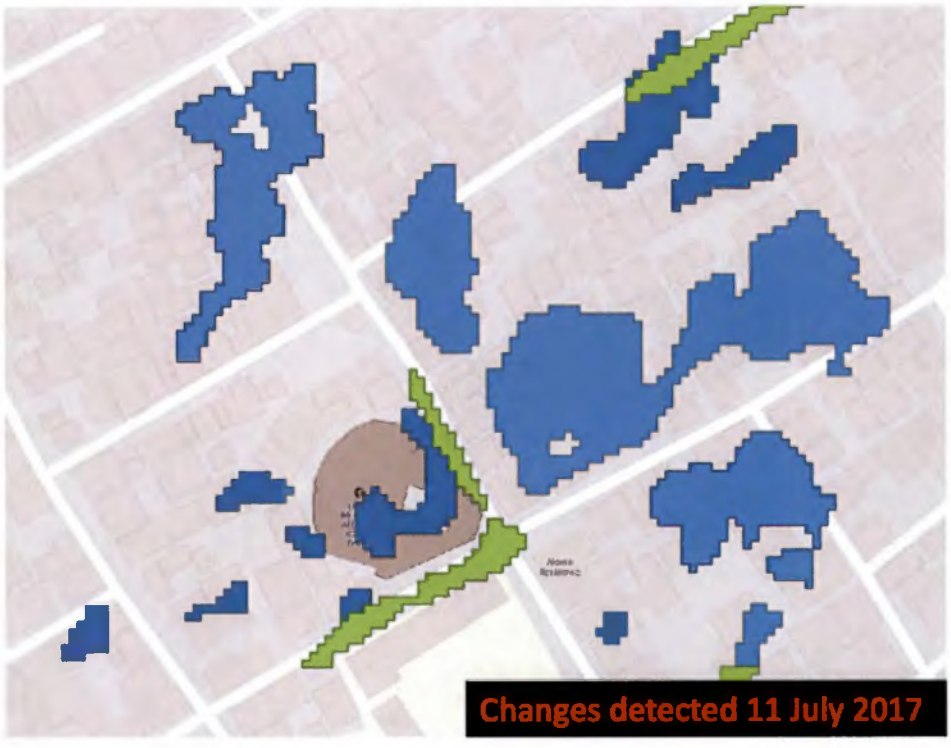

Figure 22. A close up of changes near the Masjed Mindanao Islamic Centre.

By looking at a close of the mosque area, we can see large destruction polygons (blue). These changes correspond to reports of militants using the mosque as a stronghold, and the government's operational response ("The Straits Times," 2017 August, 26). The changes were detected using a land cover change detection script following the mass-classification of scenes using the land cover algorithm. This algorithm was trained on Planet Dove imagery with a 3m ground sampled distance (GSD) resolution. Additionally, the construction polygons of roads (green), show how

changes in the road are portrayed over time. Imagery in the "before" time range most likely had the road covered with debris or barricaded, and therefore that area would not be classified as road. It would make sense that rubble would be routinely created by bombings, and then removed to some degree in order to facilitate mounted patrols or other vehicle-necessary military operations. In a non-permissible environment such as a battle or humanitarian crisis, ground truth is often difficult to attain. Classifying large areas at scale helps provide a user of this technology make better decisions about what is happening on the ground as opposed to them not having any data and trying to operate blindly with no guidance. Figure 23 shows a higher resolution image from Digital Globe at the same location as the land cover classification outputs in figure 22. Using a lower resolution imagery source with a higher revisit rate can help detect changes, which an analyst can follow up on in higher resolution imagery for a closer analysis.

Figure 23. High resolution Digital Globe image taken a day after the land cover algorithm detected change in Planet imagery.

Even amongst different commercial imagery sources, a variation of "tip and cue" could be used by one user or analyst with access to the algorithms. The tip and cue concept is used as a solution to the problem of trying to sort through loads of data efficiently to identify key events and patterns of life. Tipping takes place when key characteristics in an image in a large database is discovered and alerts an analyst or other user, thus saving that analyst from searching for the events in the database blindly (Post, 2017). Cueing then is the creation of a list of characteristics to identify in the image database. Often the tipping event is generated after analyzing a large amount of imagery, and the cueing component allows for an analyst or other algorithm to take a closer look with higher resolution imagery. It often is a use case for sensors that are over tasked or have limited resources (Post, 2017).

Figure 24 shows four different change detection results from four different time periods. The overall time frame was June 2017 to November 2017. Baseline and target time ranges were used on a month to month basis. The green polygons indicate the construction of structures just outside the main battle area along a main highway. Upon seeing this, an analyst would want to investigate further.

Figure 24. Construction polygons from different time ranges as the battle progresses, left to right.

The land cover change detection was the tip, and the cue would be a follow up to get more details with higher resolution imagery. In this case, Digital Globe Worldview imagery was used in the follow up (figure 25). The outcome showed a field hospital and helicopter landing pad construction. This finding shows how automation helps consumers of large quantities of remotely sensed data find value quickly, which in turn is implemented into mission or other plans.

Figure 25. Digital Globe Worldview high resolution imagery of figure 21. The helipad and construction of new buildings could indicate a field hospital, or other AFP/humanitarian installation. By identifying other

such installations, a pattern of life can be pieced together. Detecting this trend without the help of the LULC algorithm would require increased analyst time and resources.

Quantifiable metrics for building and road class area also help understand pre-, during-, and post-battle events. Not surprisingly, the number of square meters destroyed (figure 26) would most likely be more than the number of square meters constructed (figure 27) during the battle. These metrics help understand when the most buildings were destroyed during the battle on a high level. This information is important for post battle analysis as well as near-real time metrics as the battle is going on. For reconstruction and humanitarian efforts, it helps quantify square meters of buildings or roads constructed post battle, thereby tracking reconstruction progress quantifiably.

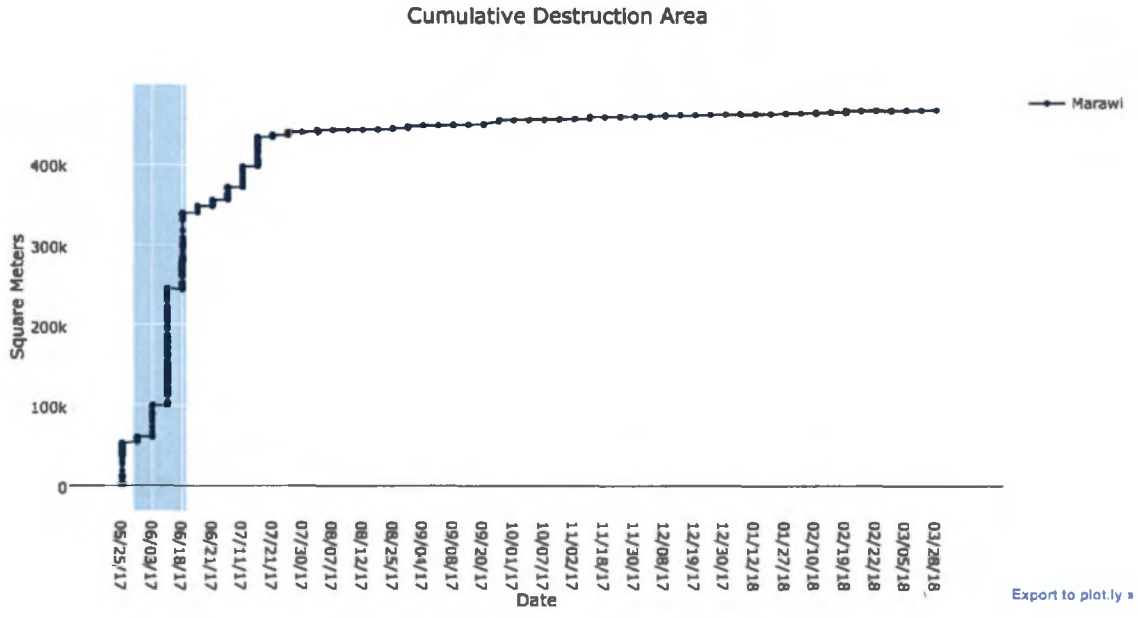

Figure 26. Square meters of area destroyed calculated based off land cover algorithm classifications. Shaded blue area indicates duration of battle. AOI is the greater Marawi area. Made with Plotly.py library.

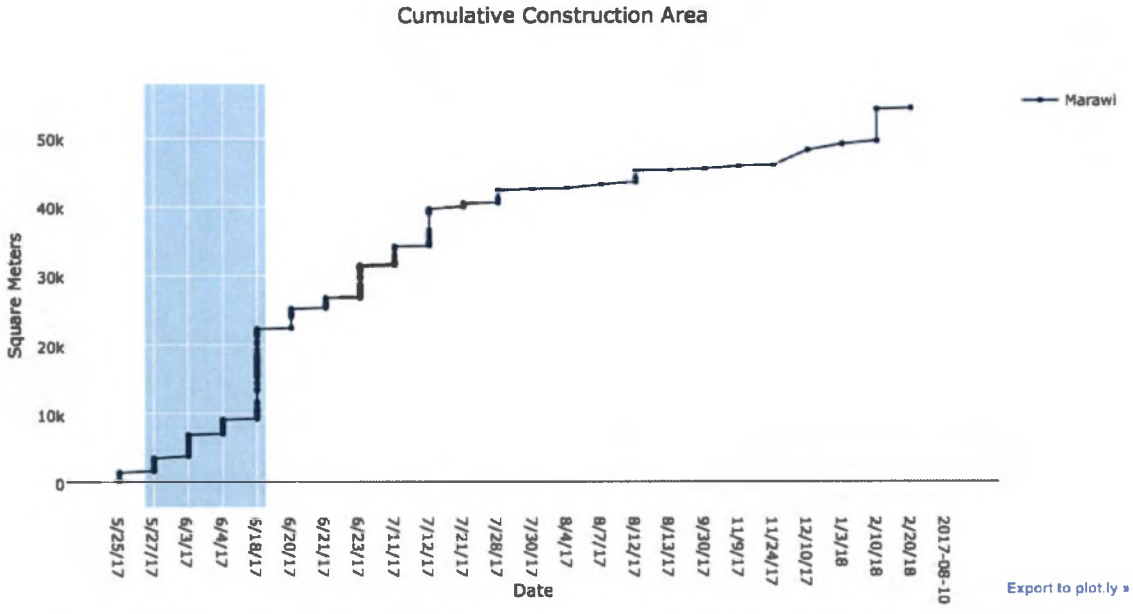

Figure 27. Square meters of area constructed calculated based off land cover algorithm classifications. Shaded blue area indicates duration of battle. AOI is greater the Marawi area. Made with Plotly.py library.

The main land cover analysis can be summarized by the following. First, automatically derived class changes at a large scale allows an analyst to more easily detect anomalous activity. This gives the analyst a visual guide on which images to start analyzing in detail. Second, when comparing land cover output at scale before and after the battle, or month to month, the building and road classes identify areas requiring further analysis from a military and humanitarian perspective. Third, the AOI can automatically be examined over a long period of time to evaluate the speed at which reconstruction and repopulation takes place. A user or analyst can then take this data and incorporate it into their intelligence or operational models. Previously, manual gathering of this data would take additional time and resources.

2.1.4 Geolocation Data

In addition to satellite imagery, geolocation data is a relatively new form of spatial dataset that is being used for PoL analysis. The geolocation data gathered for this analysis is based on anonymous cell phone GPS. Mobile applications on cell phones sometimes require location services to be enabled. This data is aggregated and sold to third-party vendors, which in turn sell it to other consumers of data. For example, geolocation, or "telemetry cell phone data", is used to help understand traffic patterns in cities, thus making transportation more efficient (Bachir et al., 2019). The data comes in the form of "pings." Pings are when a unique device such as a cell phone transmits a signal. That signal is then cached, along with billions of other pings from other devices. The pings contain coordinates and other data that can be analyzed to understand trends of movement, dwell time, etc. Visually, pings can be thought of as a point shapefile layer.

This data has restrictions depending on the country, and the European Union (EU) has enacted the General Data Protection Regulation on 25 May 2018 (Goddard, 2017). GDPR increases the jurisdiction of those protected, the penalties of people breaching such data, strengthened conditions for consent, among other privacy based regulations regarding personal data("Key Changes with the General Data Protection Regulation – EUGDPR," n.d.). Geolocation data also has different "penetration rates" between countries, according to the third-party providers. The Philippines does not abide by GDPR and has a penetration rate high enough to understand basic insights into the Marawi area.

Geolocation data is acquired by different methods, but for our brief examination, the geolocation comes from a third-party provider that bought the data from undisclosed mobile apps. These apps were present on an individual device (mobile phone) during the time of location data retrieval. A one-month global sample was acquired from 15 June 2018 to 15 July 2018. Although the sample timeframe is not ideal regarding the time of the battle, it does showcase what possibilities there are when incorporating geolocation data. For example, geolocation data can identify a high-level pattern of life change within a city due to an acute decrease of counts. It can also help quantify displaced persons' migration patterns when only qualifiable information is available (Stein, 1981). Along with imagery, the geolocation data shows we can help identify "cold zones" or bombed out areas where people would not be living. For this use case, basic analysis concerning total number of pings and unique device counts over Marawi (figure 28), and a nearby city (figure 29) not engaged in a battle are compared. By visually looking at the data, the area with the heaviest fighting is a "cold zone." Table 2 shows the breakdown of unique device counts and total pings for each city. This is just one example showing what can be derived from geolocation data. By adding context, personnel can more specifically query locations for geolocation data, proving more information for their specific need.

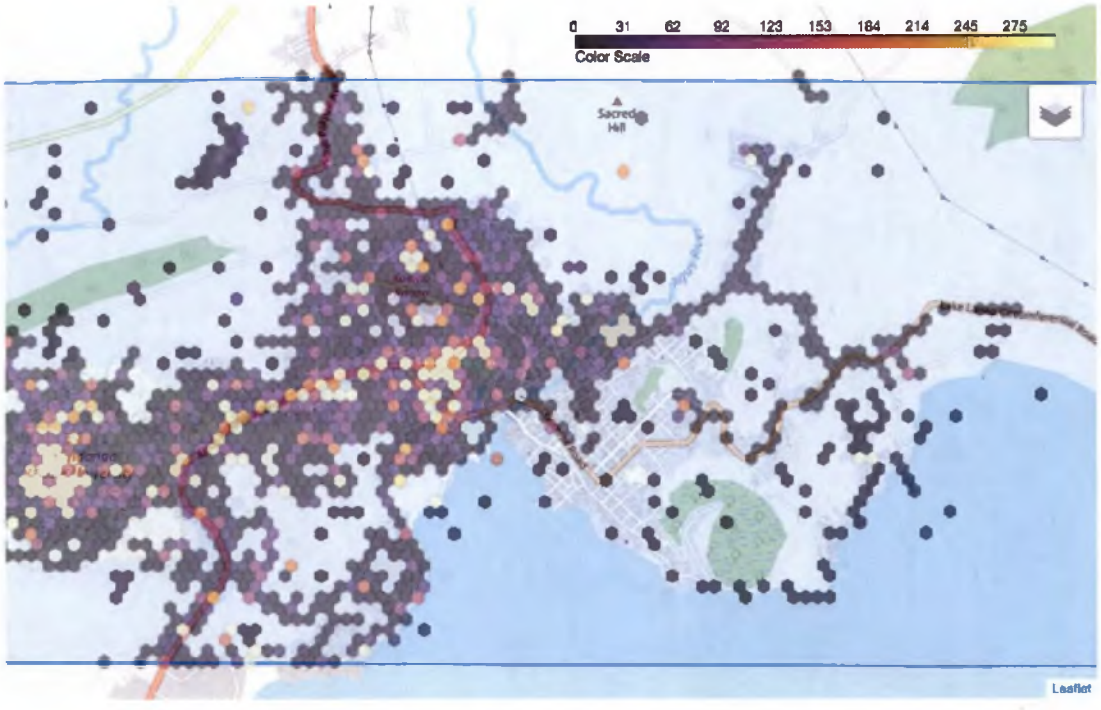

Figure 28. Geolocation hexbin plot in the greater Marawi area. Note that despite the battle being over for about 8 months, geolocation pings in the southeast part of the city are still reduced. Made with Leaflet JavaScript library in a Python Jupyter Notebook.

City	Total Number of Pings	Total Number of Unique Devices
Marawi	117,336	4,099
Iligan	461,934	18,624

Table 2. The total pings and unique device count only show part of what is happening on the ground. Visualizing clustering and dispersion are important to understand human patterns of life.

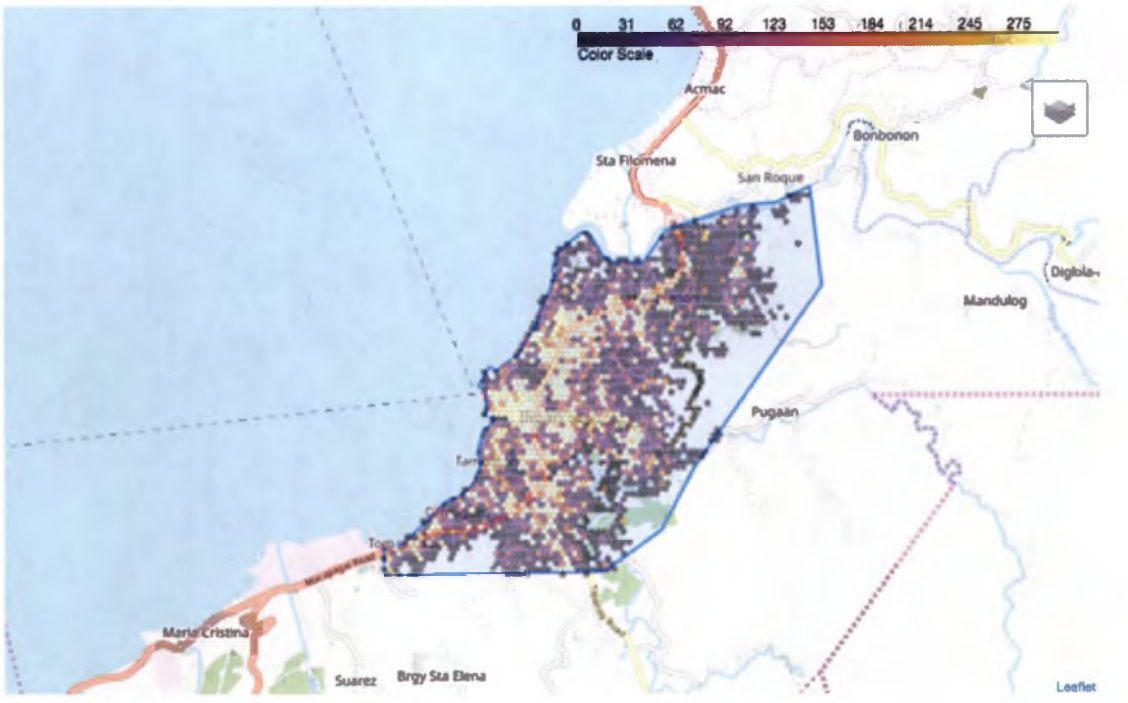

Figure 29. Geolocation hexbin plot in the greater Iligan area, a neighboring city. Notice the lack of a "cold zone," and the homogeneity of pings as one would expect in a city. Made with Leaflet JavaScript library in a Python Jupyter Notebook.

2.2 Indications and Warnings: Multi-class Aircraft

2.2.1 Introduction

Another use case revolves around high-level indications and warnings. Indications and warning intelligence is a methodology used by analysts to produce intelligence that can be used by the military or policy makers regarding an adversary's ability and likelihood to act dangerously or aggressively (Wirtz, 2013). The idea is that by monitoring certain characteristics of an enemy, operational decisions will be made early on so that preventative action can take place before any nefarious action comes to fruition (Wilson et al., 2008). To be successful at implementing policies from indication and warning analysis, data is needed to identify patterns and sift through the noise (Lasoen, 2017). This data can take many forms, but for this use case it is attempted to create a

workflow that can be implemented into a larger threat analysis model for entire countries or regions. The regions examined were military or dual-use airfields in the countries of China and Russia. Other airfields in the Middle East were also examined. By counting aircraft, policymakers can begin to understand the quantifiable counts of aircraft associated with each region or specific airfield and what is normal. These airfields can be continuously monitored, and any anomalous behavior automatically detected by the computer vision algorithm for multi-class aircraft detections. This use case also will attempt to show how new technology is able to use commercially available imagery and create historical time series analysis at scale for entire countries and their military aircraft posture.

2.2.2 Methods

The multi-class aircraft detector classifies fighter, bomber, and commercial/passenger, or other types of aircraft. Using cloud computing, 342 airfields and dual-use airports are analyzed. Historical imagery over the AOIs helps create a baseline pattern of life. As more training data is created and the algorithm is improved, new classes of aircraft can be incorporated into the multi-class detector. Large scale counts of these classes, particularly of near-peer adversaries, helps guide strategic plans such as the National Defense Strategy and National Military Strategy. The audience for this type of technology includes military analysts at all levels of command or intelligence analysts working in various organizations and teams.

More specifically, the challenge that the multiclass aircraft detector algorithm is trying to solve is understanding foreign military aircraft disposition of forces at hundreds of locations. During large-scale conflicts, monitoring kinetic enemy activity uses valuable time and resources. Automatically analyzing historical imagery produces a baseline aircraft count. The baseline provides the analyst a database to refer to that includes historical aircraft locations and counts. Anomalous counts of aircraft are then more easily

identifiable. This analysis is relevant during peacetime and combat operations of varying scales.

To showcase the importance the algorithm provides for saving time and resources, a project was created to count aircraft from available Digital Globe imagery across all 342 airfields for 2018. To further the granularity of this analysis, a user could take the aircraft classes and break down the analysis by groups of airfields or type of aircraft.

2.2.3 Results and Discussion

The multi-class aircraft algorithm counted 53,974 aircraft from 2,469 imagery scenes over 342 airfields. The algorithm ran on imagery filtered for less than 60% cloud cover with a minimum area coverage of 10%. This means that at least 10% of the image scene needed to be within the AOI (airfield) boundaries. The automatic analysis took slightly under a day to run. Conversely, it would take a human analyst a conservative and optimistic estimate of 7.5 work days to perform the same task. A breakdown of human counting estimates is provided in table 3. Human count estimates were gathered from the writer, who helped create the training data for the multi-class aircraft detecting algorithm. With computer vision and automated workflows, an analyst can focus more on contextual details of an area, providing a more accurate analyst report. The automated workflow also allows for a more rapid decision-making process during operational planning.

Number of scenes	Number of aircraft in scene	Avg. seconds it takes a human to count/analyze aircraft per scene	Seconds	Hours
517	0	5	2,585	
709	1-10	15	10,635	
332	11-20	45	14,940	
446	20-40	90	40,140	
457	40	300	137,100	
total			**205,400**	**57**

Table 3. Breakdown of estimated time it would take a human to count aircraft. It would take 57 straight hours of work to perform what the algorithm performed in slightly less than 24. This equates to approximately 7.5 work days assuming an 8-hour work day and breaks.

Large scale analysis is one key result of using cloud computing, AI/ML/CV, and improved imagery ingestion pipelines. These technologies allow for a workflow that could not be done with traditional remote sensing software. With the correct architecture, hundreds of AOIs are analyzed. Figure 30 shows the AOI locations of the airfields monitored with available Digital Globe imagery from 2018. The multi-class aircraft detector algorithm then ran on that imagery as one complete, end-to-end workflow.

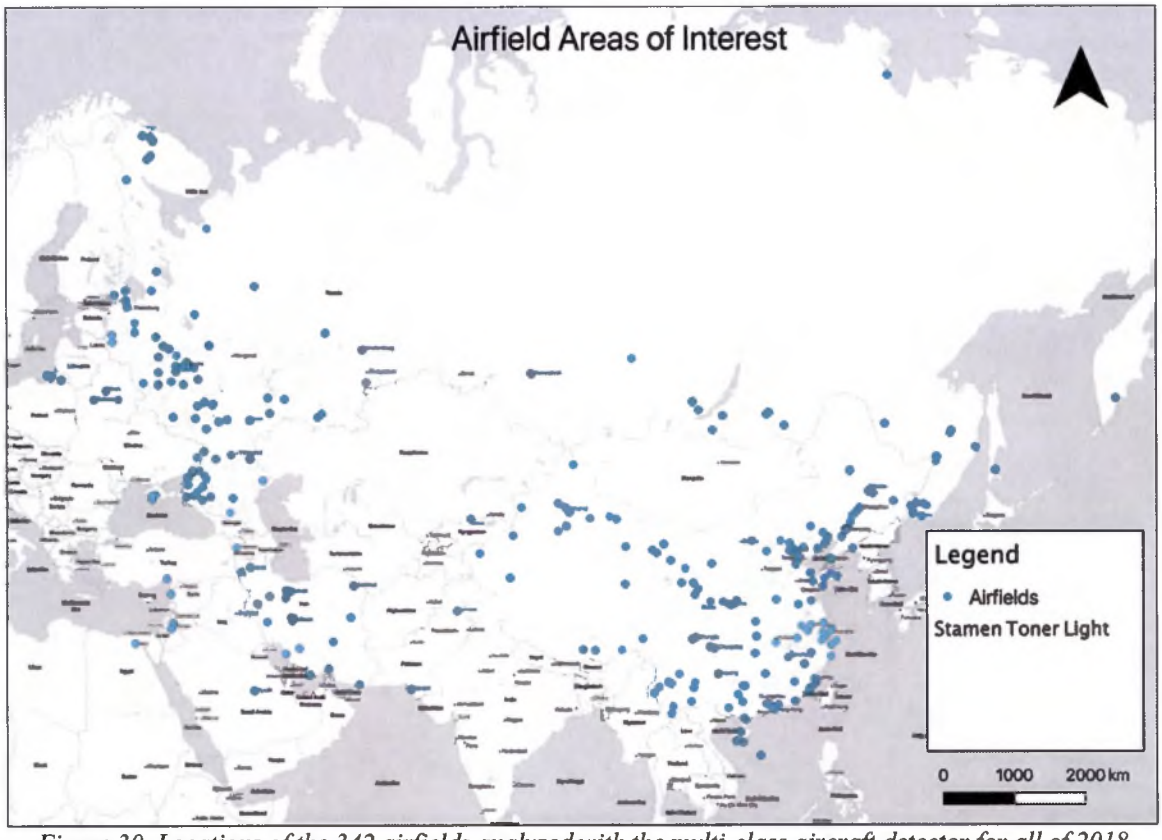

Figure 30. Locations of the 342 airfields analyzed with the multi-class aircraft detector for all of 2018 using available Digital Globe imagery. Made with QGIS v3.0

Once the algorithm is finished analyzing imagery, raw data is acquired via an application program interface (API) endpoint. Figures 31 - 34 are graphs created with the measured count of aircraft for an AOI, or groups of AOIs. The derived data shows aircraft locations as well-known-text (WKT) format. WKT is a text markup language for representing vectors on a map, and can be used to visualize points, lines, polygons, multi-point, multi-polygons, etc. ("Well-known text representation of geometry," 2019) WKTs act similarly to coordinates or shapefiles, as the text string can be used to create a visual geometry in a Geographic Information System.

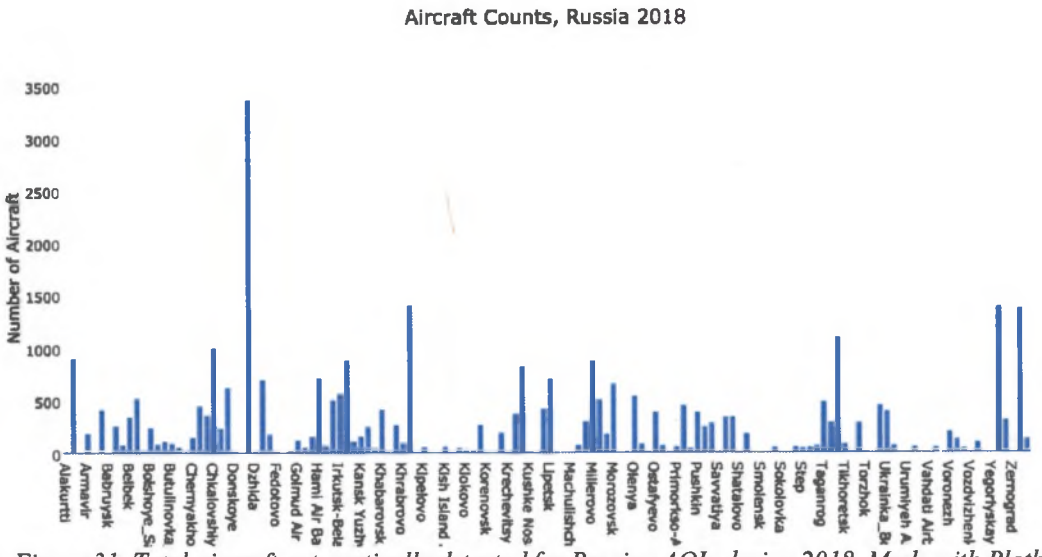

Figure 31. Total aircraft automatically detected for Russian AOIs during 2018. Made with Plotly.py library.

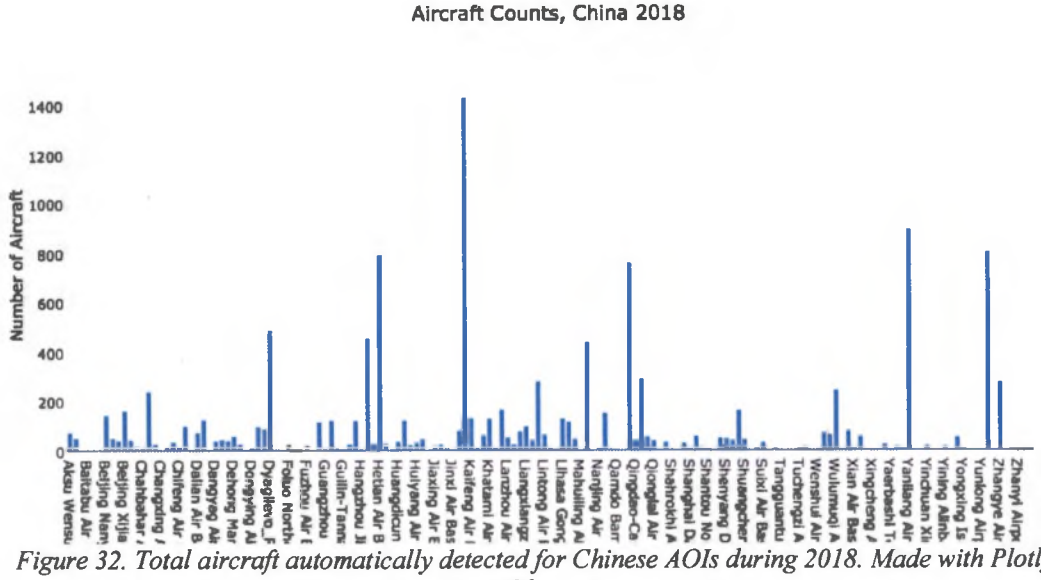

Figure 32. Total aircraft automatically detected for Chinese AOIs during 2018. Made with Plotly.py library.

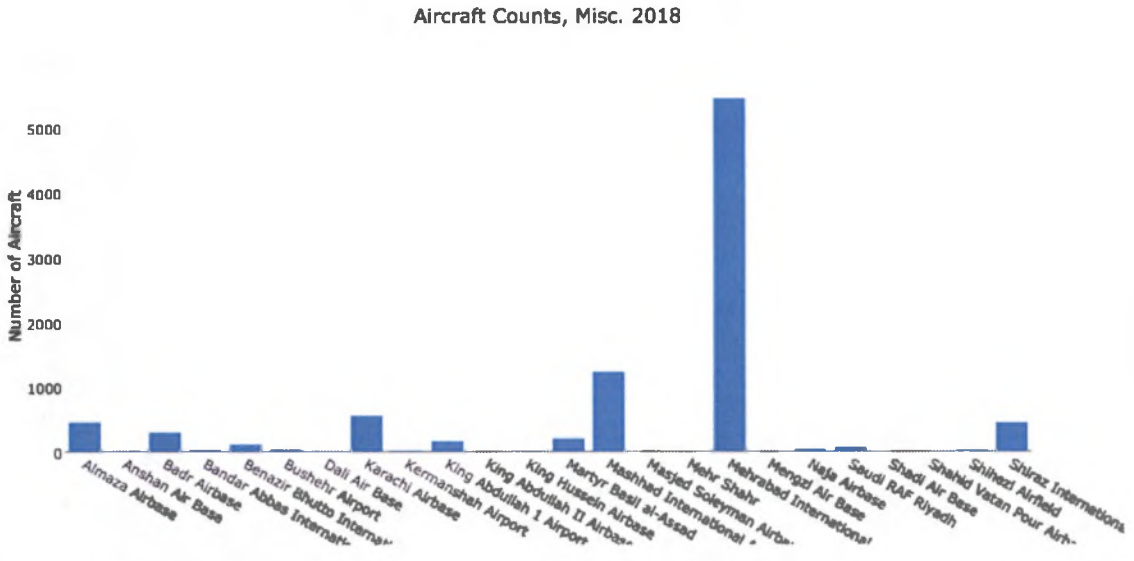

Figure 33. Total aircraft automatically detected for miscellaneous AOIs during 2018. Made with Plotly.py library.

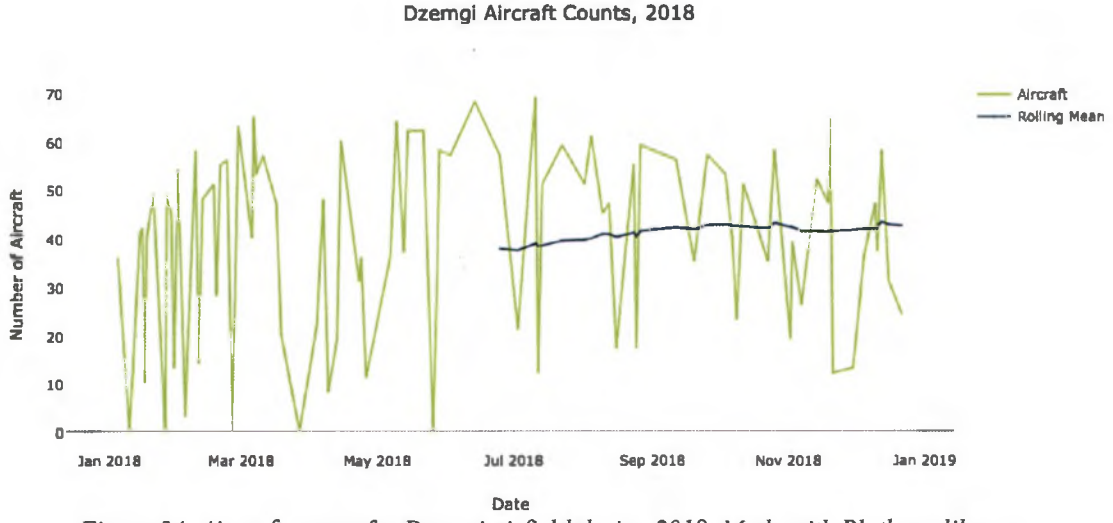

Figure 34. Aircraft counts for Dzemgi airfield during 2018. Made with Plotly.py library.

The multi-class air detector algorithm results output is in the form of a bounding box as shown in figure 35. These rectangles are categorized as fighter, bomber,

commercial/passenger, or other. Close-ups of thumbnail Digital Globe imagery in figures 36 and 37 show ground truth detections of aircraft.

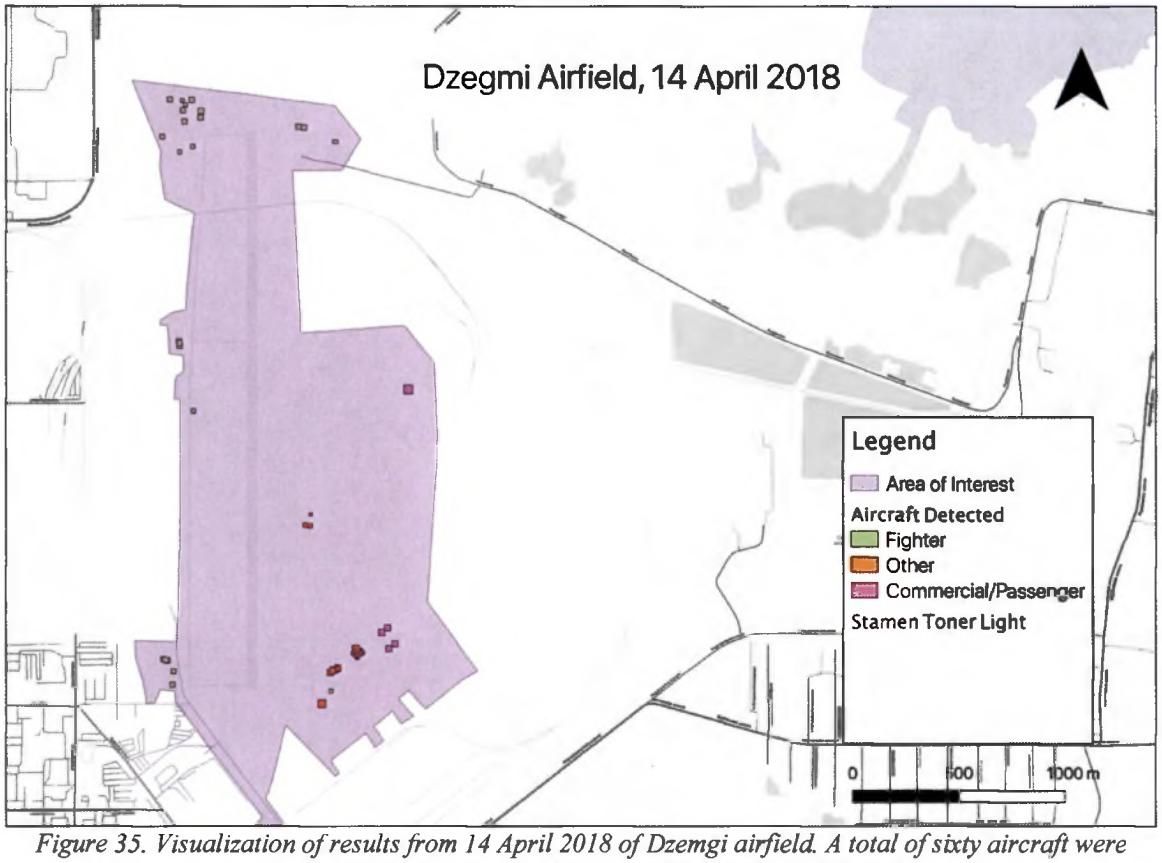

Figure 35. Visualization of results from 14 April 2018 of Dzemgi airfield. A total of sixty aircraft were automatically detected in this image. Made with QGIS v3.0

Figure 36. Close up of aircraft detected in figure 35. The image corresponds to the northwest portion of the AOI in figure 35. Screenshot taken from Orbital Insight platform using Digital Globe thumbnail as base map.

Figure 37. Close up of aircraft detected in figure 35. The image corresponds to the middle group in the southeast portion of the AOI in figure 35. Screenshot taken from Orbital Insight platform using Digital Globe thumbnail as base map.

In peacetime, monitoring indications and warnings via disposition of near-peer enemy aircraft saves time and resources. It allows the military or intelligence analyst to focus on context and create deeper analysis while the data is being automatically collected. This historical data can also help predict when anomalies happen. For example, if airfields near a contested region suddenly have an increase of fighters or bombers, it could indicate some type of pending military action. The historical data collection of those specific AOIs, and the more robust analysis from the analyst, will help verify enemy intention. This in turn helps military and intelligence strategists formulate a more rapid and accurate mission response plan. Analysis can also be broken down into regions or individual airfields, with analysts assigned specific AOIs. This top-down planning/bottom up refinement flow of data creates ownership and transparency among an organization or military unit.

2.3 Camp Wildfire

2.3.1 Introduction

The Camp Fire in Butte County, CA was the deadliest and most destructive wildfire in California history ("CAL FIRE - Home," n.d.). Total burned area is estimated to be 153,336 acres, with 86 fatalities and 18,804 buildings destroyed ("Camp Fire (2018)," 2019). There have been many studies surrounding forest recovery from wildfires using remote sensing. The use of the Normalized Difference Vegetation Index (NDVI), which is a numerical indicator that uses visible and near-infrared (NIR) bands to assess whether what is being observed contains live green vegetation or not has been used many times to evaluate forest recovery (Cuevas-González et al., 2009). The shortwave infrared band (SWIR) has also been used to identify active fires, and to some degree see through smoke (Cho et al., 2018). Despite these analyses being useful, there are limitations. First, many of the analysis being used are run on imagery that is of relatively medium to low resolution (Brewer, et al., 2005). Additionally, the higher resolution imagery coming

from airborne and drones are only able to image small parts of a large fire, and are relatively expensive to employ (Allison et al., 2016). In order to provide valuable, timely, and large-scale information during a wildfire, new methods need to be employed that take advantage of cloud computing, AI/ML/CV, and robust imagery ingestion pipelines.

2.3.2 Methods

A pipeline using cloud computing to rapidly ingest imagery allows for automated analysis to take place on imagery the same day of capture. Once imagery is ingested an already built land cover algorithm processes the imagery and provides land cover classifications over a large area. This can be repeated daily or as imagery becomes available. The area classified was 620 square kilometers. Before and after periods of one month each were used to show results, but in a real-time scenario time ranges can be adjusted as needed. The land cover aggregation algorithm used takes an area of interest (AOI) and ingests available imagery based on certain parameters such as date range, cloud cover, and percent AOI coverage. It then looks at all cells from all scenes and assigns it a class based on all the available data. These cell classifications are then polygonised. The polygon results are visualized in a web-based platform built by Orbital Insight. Additionally, the result geometries can be retrieved via an API and visualized in any GIS software.

Once the imagery request was completed, the algorithm took approximately 3 hours to complete the analysis. The result is an aggregated classification of the Camp Fire extent area. In this manner, we can collect such foundation data from time periods closely preceding a natural disaster, thereby offering a viable and accurate baseline comparison against other data sources for after the disaster. This provides government and humanitarian organizations with real-time, predictive insights regarding the human, infrastructure, and network costs of these events.

2.3.3 Results and Discussion

Visually the differences between figures 38 and 39 can be seen even when zoomed out to show the full extent of the Camp Fire's reach. Not surprisingly, large areas of the forest class were removed. These classifications, if performed daily, would be able to give a quantifiable estimate of class loss quicker and easier than a person on the ground. Once the data is in a GIS or Pandas python library DataFrame ("Python Data Analysis Library — pandas: Python Data Analysis Library," n.d.), it can be analyzed to quantify changes in classes. Additionally, one can zoom in as shown in figures 40 and 41 to see more specific instances of destruction. Figure 40 shows a neighborhood where buildings were classified in the before aggregation but lost in the after aggregation, and figure 41 shows another area affected by the Camp Fire with before and after land cover classification results. figure 42 attempts to portray visual accuracy of the algorithm. Classification results output is overlaid on an OpenStreetMaps base layer for comparison.

Areas with no classification simply did not have enough data as per the aggregation parameters of the algorithm. These thresholds can be adjusted if needed to accept a wider variance of classification precision metrics. However, with less stringent precision thresholds for each class, the higher the potential for misclassifications and therefore reduced usefulness.

Metrics concerning damage assessment are important to policy makers and leaders in government, even while a fire is ongoing. Anecdotally, firefighters are often ordered to count damaged and destroyed buildings even when the extent of the fire is unknown. Due to these politics, government leaders succeed in filing reports but at the expense of firefighter resources that could better have been spent by obtaining where the fire is in entirety and fight it appropriately.

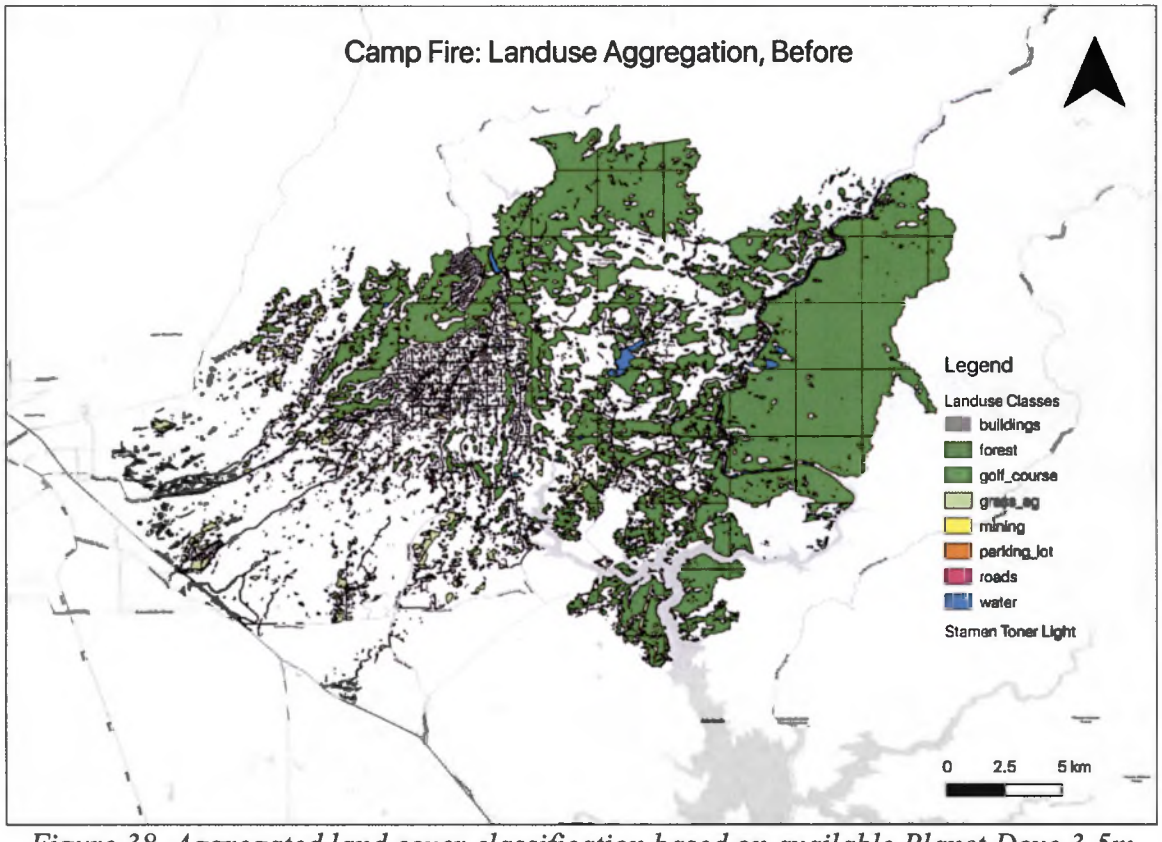

Figure 38. Aggregated land cover classification based on available Planet Dove 3-5m GSD imagery for one month prior to the start of the Camp Fire. Made with QGIS v3.0.

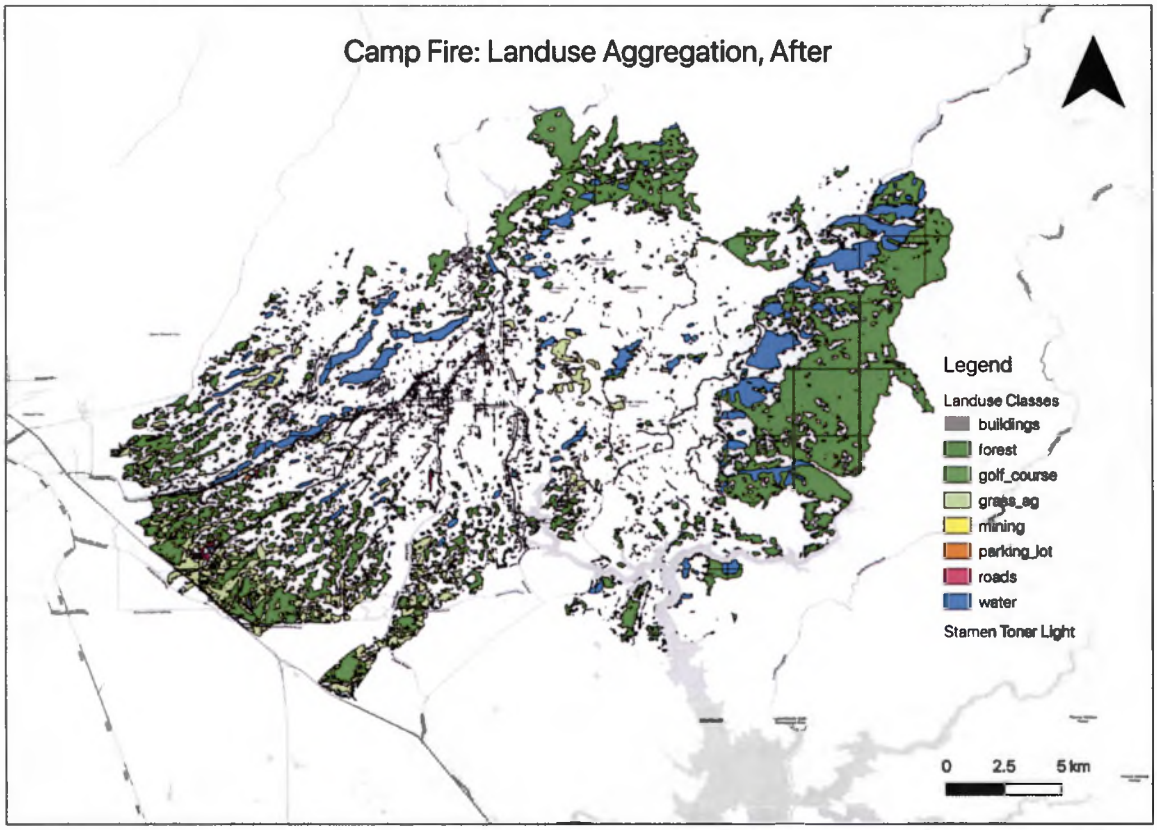

Figure 39. Aggregated land cover classification based on available Planet Dove 3-5m GSD imagery for one month after the end of the Camp Fire. Note: false positives for the water class , but accurate portrayal of forest and building class removal. Made with QGIS v3.0.

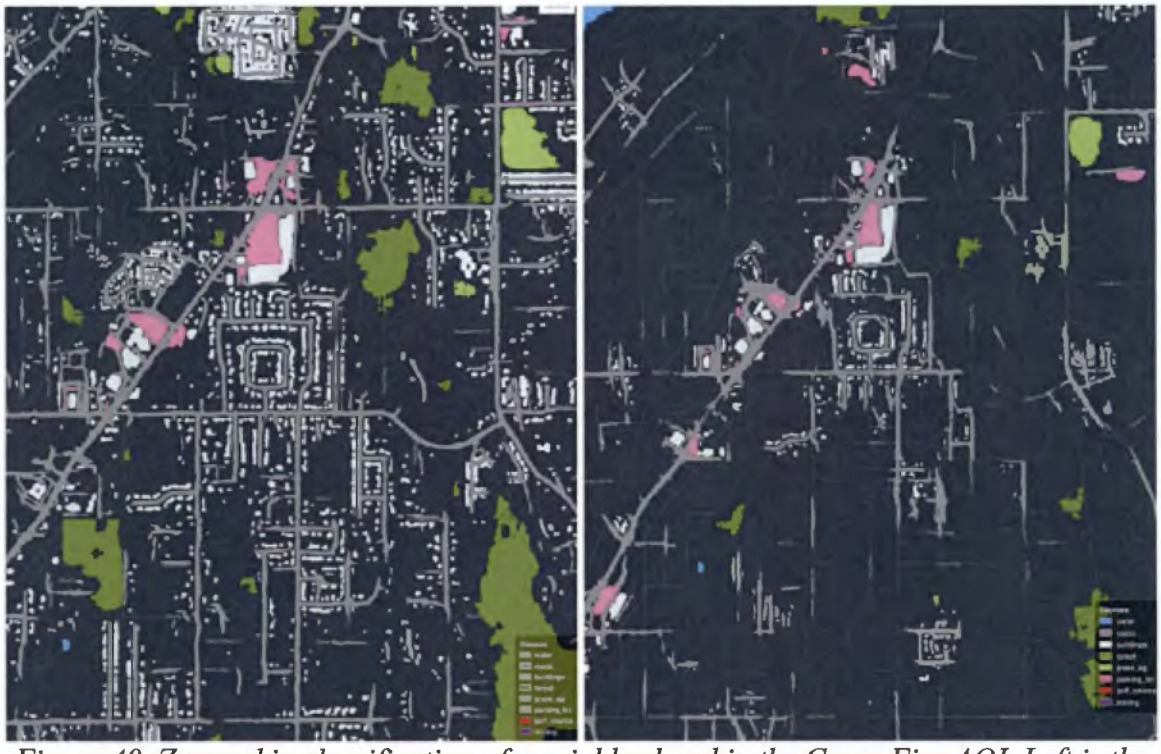

Figure 40. Zoomed in classification of a neighborhood in the Camp Fire AOI. Left is the before fire classification result. Right is the after-fire classification result. Light gray is buildings, dark gray is roads, dark green is forest. Screenshot taken from Orbital Insight platform using Mapbox basemap.

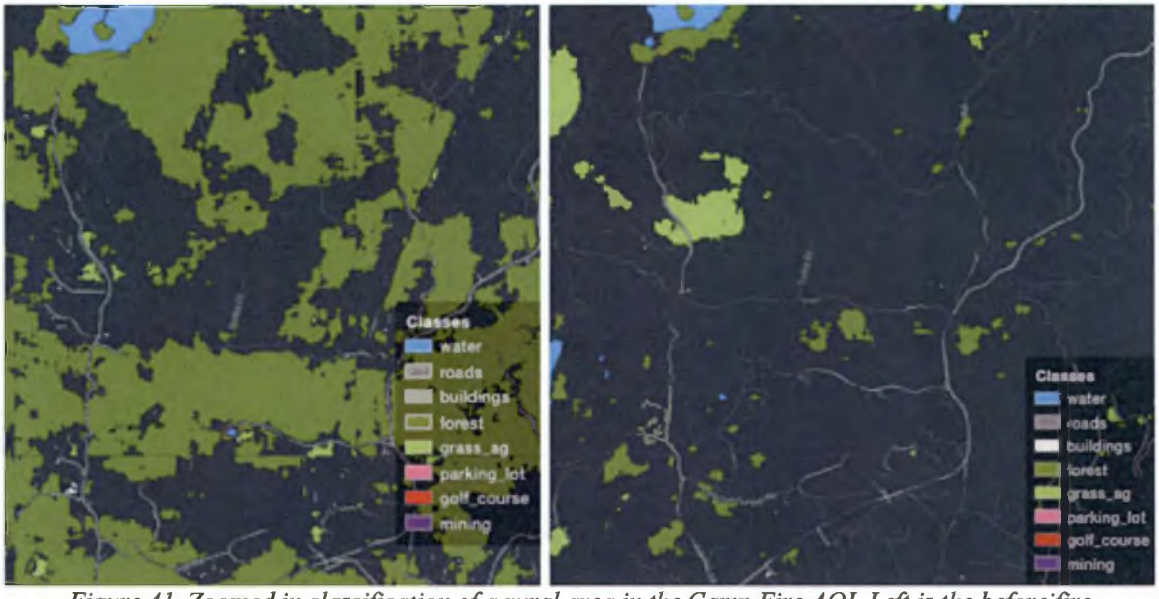

Figure 41. Zoomed in classification of a rural area in the Camp Fire AOI. Left is the before-fire classification result. Right is the after-fire classification result. Light green is grass, dark gray is roads, dark green is forest. Screenshot taken from Orbital Insight platform using Mapbox basemap.

Figure 42. Close-up comparison of land cover aggregation algorithm and OpenStreetMaps basemap, before fire, showing building (red) and road (pink) classes. Made by using QGIS v3.0.

Conclusion

This study showcased the advantages cloud computing, AI/ML/CV, and big data analysis has on remotely sensed data for humanitarian and defense issues. Three use cases were examined to illustrate how workflows using these technologies assist defense and humanitarian organizations achieve mission successes.

The Battle of Marawi used object detection, land cover change detection, and geolocation data to show how a commander on the ground could better fill in information

gaps to better create an operational plan. Humanitarian organizations could use the data to better project civilian pattern of life migration and rates of city reconstruction.

An indications and warnings use case used a multi-class aircraft detecting algorithm as a proxy for near-peer adversary military activity. Aircraft counts were taken form 342 airfields over a one-year time period. This data helps create a baseline, which can be incorporated into a military or intelligence model for high-level foreign operational activity. It also allows a user to create a historical analysis from which future anomalous activity can be automatically detected, saving analyst time and resources.

Last, it was briefly examined how with the proper imagery ingestion pipelines and automated analysis, an algorithm can mass-classify large disaster areas before and after the event. This analysis helps with post-disaster reporting as well as relatively real time data needed to guide disaster relief efforts. The algorithm can provide daily metrics on land cover classes that otherwise would have needed a human to manually review all available imagery. This in turn allows more firefighters and relief personnel to be in the area where action is needed most, and not wasting time counting buildings that have been damaged, something an algorithm can instead do.

The use cases show that as cloud computing, AI/ML/CV, and remotely sensed data become more mainstream in private and public sectors, there are manifold possibilities it can be used to help solve humanitarian and defense problems. The algorithms vary in maturity, iterations built, number of training datasets, and imagery used. As algorithm performance increases, so will their use by humanitarian and defense organizations.

Algorithm performance proved well enough to understand use cases better and save a user time and resources. However, these algorithms need constant training and increased iterations that encompass more diverse training sets. Additionally, industry would find it useful to train algorithms on open source imagery for humanitarian or other public sector use in local, state, and federal governments. Having these assets ahead of a

humanitarian crisis will help officials be prepared in correctly allocating resources to the correct areas in an efficient manner.

References

Allison, R. S., Johnston, J. M., Craig, G., & Jennings, S. (2016). Airborne Optical and Thermal Remote Sensing for Wildfire Detection and Monitoring. *Sensors*, *16*(8), 1310. https://doi.org/10.3390/s16081310

Amazon EC2 Instance Types - Amazon Web Services. (n.d.). Retrieved May 1, 2019, from Amazon Web Services, Inc. website: https://aws.amazon.com/ec2/instance-types/

AWS Ground Station – Ingest and Process Data from Orbiting Satellites. (2018, November 27). Retrieved April 1, 2019, from Amazon Web Services website: https://aws.amazon.com/blogs/aws/aws-ground-station-ingest-and-process-data-from-orbiting-satellites/

Bachir, D., Khodabandelou, G., Gauthier, V., El Yacoubi, M., & Puchinger, J. (2019). Inferring dynamic origin-destination flows by transport mode using mobile phone data. *Transportation Research Part C: Emerging Technologies*, *101*, 254–275. https://doi.org/10.1016/j.trc.2019.02.013

Belward, A. S., & Skøien, J. O. (2015). Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS Journal of Photogrammetry and Remote Sensing*, *103*, 115–128. https://doi.org/10.1016/j.isprsjprs.2014.03.009

Boschetti, L., Flasse, S. P., & Brivio, P. A. (2004). Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto Boundary. *Remote Sensing of Environment*, *91*(3–4), 280–292. https://doi.org/10.1016/j.rse.2004.02.015

Brewer, C. K., Winne, J. C., Redmond, R. L., Opitz, D. W., & Mangrich, M. V. (2005, November). Classifying and Mapping Wildfire Severity [Text]. https://doi.org/info:doi/10.14358/PERS.71.11.1311

Buchen, E. (2015). Small Satellite Market Observations. *AIAA/USU Conference on Small Satellites*. Retrieved from https://digitalcommons.usu.edu/smallsat/2015/all2015/51

Buonaiuto, N., Louie, M., Aarestad, J., Mital, R., Mateik, D., Sivilli, R., … Zufelt, B. (2017). Satellite Identification Imaging for Small Satellites Using NVIDIA. *AIAA/USU Conference on Small Satellites*. Retrieved from https://digitalcommons.usu.edu/smallsat/2017/Alternates/5

CAL FIRE - Home. (n.d.). Retrieved May 10, 2017, from http://calfire.ca.gov/serp?q=lookouts&cx=001779225245372747843%3Ableoj5gjmwq&cof=FORID%3A10&ie=UTF-8

Camp Fire (2018). (2019). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Camp_Fire_(2018)&oldid=891266277

Chi, M., Plaza, A., Benediktsson, J. A., Sun, Z., Shen, J., & Zhu, Y. (2016). Big Data for Remote Sensing: Challenges and Opportunities. *Proceedings of the IEEE*, *104*(11), 2207–2219. https://doi.org/10.1109/JPROC.2016.2598228

Chisuwo, C. (2018, October 4). The eyes in the sky: On satellite imagery as big data. Retrieved March 31, 2019, from Towards Data Science website: https://towardsdatascience.com/the-eyes-in-the-sky-on-satellite-imagery-as-big-data-210830722f3d

Cho, K., Kim, Y., & Kim, Y. (2018). Disaggregation of Landsat-8 Thermal Data Using Guided SWIR Imagery on the Scene of a Wildfire. *Remote Sensing*, *10*(1), 105. https://doi.org/10.3390/rs10010105

Cuevas-González, M., Gerard, F., Balzter, H., & Riaño, D. (2009). Analysing forest recovery after wildfire disturbance in boreal Siberia using remotely sensed vegetation indices. *Global Change Biology*, *15*(3), 561–577. https://doi.org/10.1111/j.1365-2486.2008.01784.x

EDT, G. B. O. 9/8/16 at 7:10 A. (2016, September 8). Satellite imaging will revolutionize the way information is gathered, parsed and monetized. Retrieved March 31, 2019, from Newsweek website: https://www.newsweek.com/2016/09/16/why-satellite-imaging-next-big-thing-496443.html

eoPortal - Earth Observation Directory & News. (n.d.). Retrieved March 31, 2019, from https://directory.eoportal.org/web/eoportal/home

Franco, J. (2017). *The Maute Group: New Vanguard of IS in Southeast Asia?* Retrieved from https://dr.ntu.edu.sg/handle/10220/42544

Goddard, M. (2017). The EU General Data Protection Regulation (GDPR): European Regulation that has a Global Impact. *International Journal of Market Research*, *59*(6), 703–705. https://doi.org/10.2501/IJMR-2017-050

Govt troops retake grand mosque in Marawi, Asia News & Top Stories - The Straits Times. (n.d.). Retrieved April 3, 2019, from https://www.straitstimes.com/asia/govt-troops-retake-grand-mosque-in-marawi

Gunaratna, R. (2017). The Siege of Marawi: A Game Changer in Terrorism in Asia. *Counter Terrorist Trends and Analyses*, *9*(7). Retrieved from http://www.jstor.org/stable/26351533

Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Ullah Khan, S. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems*, *47*, 98–115. https://doi.org/10.1016/j.is.2014.07.006

Hassan, M. M., Smith, A. C., Walker, K., Rahman, M. K., & Southworth, J. (2018). Rohingya Refugee Crisis and Forest Cover Change in Teknaf, Bangladesh. *Remote Sensing*, *10*(5), 689. https://doi.org/10.3390/rs10050689

Hope, B. (2016, August 14). Tiny Satellites: The Latest Innovation Hedge Funds Are Using to Get a Leg Up. *Wall Street Journal*. Retrieved from https://www.wsj.com/articles/satellites-hedge-funds-eye-in-the-sky-1471207062

Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D. (2013). Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, *80*, 91–106. https://doi.org/10.1016/j.isprsjprs.2013.03.006

India space launch: One rocket, 29 satellites, three orbits. (n.d.). Retrieved April 1, 2019, from https://www.aljazeera.com/amp/news/2019/04/india-space-launch-rocket-29-satellites-orbits-190401093645706.html

Jermain, C. L., Rowlands, G. E., Buhrman, R. A., & Ralph, D. C. (2016). GPU-accelerated micromagnetic simulations using cloud computing. *Journal of Magnetism and Magnetic Materials*, *401*, 320–322. https://doi.org/10.1016/j.jmmm.2015.10.054

Key Changes with the General Data Protection Regulation – EUGDPR. (n.d.). Retrieved April 4, 2019, from https://eugdpr.org/the-regulation/

Koehrsen, W. (2018, March 3). Beyond Accuracy: Precision and Recall. Retrieved April 1, 2019, from Towards Data Science website: https://towardsdatascience.com/beyond-accuracy-precision-and-recall-3da06bea9f6c

Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker, A. L. (2016). Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, *7*(1), 3–10. https://doi.org/10.1016/j.gsf.2015.07.003

Lasoen, K. L. (2017). Indications and warning in Belgium: Brussels is not Delphi. *Journal of Strategic Studies*, *40*(7), 927–962. https://doi.org/10.1080/01402390.2017.1288111

Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017). Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(2), 645–657. https://doi.org/10.1109/TGRS.2016.2612821

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity.* Retrieved from http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation

Marr, B. (n.d.-a). How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read. Retrieved March 31, 2019, from Forbes website: https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/

Marr, B. (n.d.-b). Why Space Data Is The New Big Data. Retrieved March 31, 2019, from Forbes website: https://www.forbes.com/sites/bernardmarr/2017/10/19/why-space-data-is-the-new-big-data/

Orbital Insight | Geospatial Analytics for an Interconnected World. (n.d.). Retrieved May 2, 2019, from Orbital Insight website: https://orbitalinsight.com/

Peng, J. (n.d.). The Rise of Commercial Satellite Imaging – Implications for Defense. Retrieved April 7, 2019, from NAOC website: http://natoassociation.ca/the-rise-of-commercial-satellite-imaging-implications-for-defense/

Plotly Technologies Inc. Collaborative data science Publisher: Plotly Technologies Inc. Place of publication: Montréal, QC Date of publication: 2015 URL: https://plot.ly

Post, C. R. (2017). *Towards Automation of Tipping and Cueing Between Small Satellites in a Constellation* (No. AFIT-ENG-MS-17-M-061). Retrieved from AIR FORCE INSTITUTE OF TECHNOLOGY WRIGHT-PATTERSON AFB OH WRIGHT-PATTERSON AFB United States website: https://apps.dtic.mil/docs/citations/AD1054738

Python Data Analysis Library — pandas: Python Data Analysis Library. (n.d.). Retrieved May 1, 2019, from https://pandas.pydata.org/

QGIS Development Team (2019). QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org

Quinn John A., Nyhan Marguerite M., Navarro Celia, Coluccia Davide, Bromley Lars, & Luengo-Oroz Miguel. (2018). Humanitarian applications of machine learning with remote-sensing data: review and case study in refugee settlement mapping. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and*

Engineering Sciences, *376*(2128), 20170363. https://doi.org/10.1098/rsta.2017.0363

Refugees, U. N. H. C. for. (n.d.). A year after assault on Marawi City, residents still unable to return. Retrieved April 2, 2019, from UNHCR website: https://www.unhcr.org/news/latest/2018/5/5b0583c64/year-assault-marawi-city-residents-still-unable-return.html

Satellite Launches to Increase Threefold Over the Next Decade - Via Satellite -. (2017, October 12). Retrieved May 1, 2019, from Via Satellite website: https://www.satellitetoday.com/innovation/2017/10/12/satellite-launches-increase-threefold-next-decade/

Satellite Missions Directory - Earth Observation Missions - eoPortal. (n.d.). Retrieved March 31, 2019, from https://directory.eoportal.org/web/eoportal/satellite-missions

Sharma, B., Thulasiram, R. K., Thulasiraman, P., Garg, S. K., & Buyya, R. (2012). Pricing Cloud Compute Commodities: A Novel Financial Economic Model. *Proceedings of the 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (Ccgrid 2012)*, 451–457. https://doi.org/10.1109/CCGrid.2012.126

Singh, J. (2018). *One Year After Marawi : Has The Threat Gone?* Retrieved from https://dr.ntu.edu.sg/handle/10220/44949

Spaceflight, H. W. 2 days ago. (n.d.). Space Calendar 2019: Launches, Sky Events & More. Retrieved April 7, 2019, from Space.com website: https://www.space.com/32286-space-calendar.html

SpaceX Readies First Falcon Heavy Launch for Paying Customer - Bloomberg. (n.d.). Retrieved May 1, 2019, from https://www.bloomberg.com/news/articles/2019-04-10/spacex-s-readies-first-falcon-heavy-launch-for-paying-customer

Stein, B. N. (1981). Part VI: Resources: The Refugee Experience: Defining the Parameters of a Field of Study. *International Migration Review*, *15*(1–2), 320–330. https://doi.org/10.1177/0197918381015001-230

Tagoranao, M. S., & Gamon, A. D. (2017). The post-war reconstruction of waqf properties in Marawi City: Prospect and challenges. In *Contemporary issues on zakat, waqf and Islamic philanthrophy* (pp. 348–357). Retrieved from http://irep.iium.edu.my/60729/

Tang, W., & Feng, W. (2017). Parallel map projection of vector-based big spatial data: Coupling cloud computing with graphics processing units. *Computers, Environment and Urban Systems*, *61*, 187–197. https://doi.org/10.1016/j.compenvurbsys.2014.01.001

Tewkesbury, A. P., Comber, A. J., Tate, N. J., Lamb, A., & Fisher, P. F. (2015). A critical synthesis of remotely sensed optical image change detection techniques.

Remote Sensing of Environment, *160*, 1–14. https://doi.org/10.1016/j.rse.2015.01.006

The Most Active Space Tech Investors. (2017, May 2). Retrieved May 1, 2019, from CB Insights Research website: https://www.cbinsights.com/research/top-space-tech-startup-investors/

Weinzierl, H. (n.d.). *The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East - United States*. 7.

Well-known text representation of geometry. (2019). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Well-known_text_representation_of_geometry&oldid=881684302

Wilson, J. M., Polyak, M. G., Blake, J. W., & Collmann, J. (2008). A Heuristic Indication and Warning Staging Model for Detection and Assessment of Biological Events. *Journal of the American Medical Informatics Association*, *15*(2), 158–171. https://doi.org/10.1197/jamia.M2558

Wirtz, J. J. (2013). Indications and Warning in an Age of Uncertainty. *International Journal of Intelligence and CounterIntelligence*, *26*(3), 550–562. https://doi.org/10.1080/08850607.2013.780558