

How to Implement Drones and Machine Learning to Reduce Time, Costs, and Dangers Associated with Landmine Detection

By Jasper Baur,^{i,ii*} Gabriel Steinberg,^{i*} Alex Nikulin, Ph.D.,ⁱ Kenneth Chiu, Ph.D.,ⁱ and Timothy S. de Smet, Ph.D.ⁱ

i [Binghamton University - The State University of New York]

ii [Columbia University - Lamont-Doherty Earth Observatory]

* Equal contribution

Two rapidly emerging technologies revolutionizing scientific problem solving are unmanned aerial systems (UAS), commonly referred to as drones, and deep learning algorithms.¹ Our study combines these two technologies to provide a powerful auxiliary tool for scatterable landmine detection. These munitions are traditionally challenging for clearance operations due to their wide area of impact upon deployment, small size, and random minefield orientation. Our past work focused on developing a reliable UAS capable of detecting and identifying individual elements of PFM-1 minefields to rapidly assess wide areas for landmine contamination, minefield orientation, and possible minefield overlap. In our most recent proof-of-concept study we designed and deployed a machine learning workflow involving a region-based convolutional neural network (R-CNN) to automate the detection and classification process, achieving a 71.5% rate of successful detection.² In subsequent trials, we expanded our dataset and improved the accuracy of the CNN to detect PFM-1 anti-personnel mines from visual (RGB) UAS-based imagery to 91.8%. In this paper, we intend to familiarize the demining community with the strengths and limitations of UAS and machine learning and suggest a fit of this technology as a key auxiliary *first look* area reduction technique in humanitarian demining operations. As part of this effort, we seek to provide detailed guidance on how to implement this technique for non-technical survey (NTS) support and area reduction of confirmed and suspected hazardous areas with minimal resources and funding.

Introduction

Explosive remnants of war (ERW), including unexploded ordnance (UXO) and landmines, resulted in a recorded 5,554 casualties in 2019 with nearly half (43%) of the civilian victims—for whom the age was known—being children.³ In 2019, an estimated 164,000 emplaced anti-personnel mines were destroyed globally, but completely clearing the world of anti-personnel mines is still decades away as there are tens of millions of mines estimated to reside in place worldwide, assuming no new landmines are deployed.⁴ As of 2007, experts estimated that ten to twenty landmines were laid for every mine cleared, considerably exacerbating the landmine crisis.⁵ However, since the recent success of the implementation of the *Anti-Personnel Mine Ban Convention* (APMBC), this number may be less but is unknown. In the last decade, anti-personnel landmines have been used in active conflicts in at least fifteen countries, including but likely not limited to Afghanistan, Colombia,

India, Iraq, Israel, Libya, Myanmar, Nigeria, North Korea, Pakistan, Syria, Thailand, Tunisia, Ukraine, and Yemen.⁶ The pace of mine clearance is largely driven by operator experience, the technological capacity of mine-detection technology, and environmental difficulty, which can be quantified by survey area and ratio of successful detection relative to false flags. Recent studies demonstrated that preliminary rapid wide-area surveys conducted by unmanned aerial systems (UAS) may be utilized to initially constrain search areas, ultimately decreasing the time and cost associated with humanitarian mine action (HMA) while reducing the safety risk to clearance operators.⁷⁻¹²

Terrestrial electromagnetic-induction (EMI) methods are currently one of the main standard approaches to HMA.¹³ While hand-held EMI detector surveys have proven themselves as one of the most reliable geophysical techniques for HMA, their implementation



Figure 1. Diagram of individual PFM-1 anti-personnel landmine alongside KSF-1 dispersal cassette (adapted from de Smet et al. 2018).
All figures courtesy of the authors.

has a number of shortcomings: (1) they have high false-positive alarm rates in the presence of metallic clutter; (2) they are time and labor intensive, especially in difficult terrain; and (3) they entail operator risk. Some of these shortcomings can be mitigated if suspected hazardous areas (SHAs) are constrained by autonomous surveys providing an initial assessment of mine presence and subsequently expanded to classify mine type, mine condition, minefield density and orientation, soil type, and other environmental parameters. Modern unpiloted aerial vehicle (UAV) platforms capable of flying at low altitudes and collecting autonomous surveys, coupled with sensitive and compact visual, thermal, multispectral, and magnetic sensors, often allow operators to rapidly identify small anthropogenic targets previously identifiable only in ground surveys in certain situations.¹⁴

Over the last decade UAVs have become more reliable, and their push into the consumer market considerably decreased their costs. In parallel with the advancement of UAV technology, modern miniaturized optical and geophysical sensors became smaller, more sensitive, less costly, and mountable on UAV platforms. Advances in UAV and sensor technologies enabled the development of reliable UAS for wide-area, high-resolution remote sensing and geophysical surveys to address some of the most pressing humanitarian challenges.^{8–12} With recent developments in small autonomous UAVs, advanced sensors have the potential to significantly contribute to the field of HMA, as this allows for rapid low-cost data acquisition over wide areas in a safe and time-efficient manner. However, the emergence of UAS surveying has led to the new problem of analyzing these large (both in terms of area covered and file size) and prohibitively complex datasets, requiring advances in machine learning to aid interpretation. As we have learned firsthand, manual analysis of these large surveys is operationally difficult, subjective, and sometimes inconsistent. The application of deep learning to remotely collect wide-area surveys (greater than 15 m²) improves the reliability of NTS and provides stakeholders with the quantitative data necessary to plan HMA activities. This area reduction methodology may ultimately be used to help guide ground demining activities to reduce search area size and drive down HMA costs while reducing operational risks.

We present a machine learning case study focused on initial detection and identification of the widely-used PFM-1 anti-personnel

landmine (also known as the *butterfly mine*) as an index example of a small, low-metallic scatterable landmine. This type of mine is emblematic of a wide-area aerial-mining strategy responsible for both a legacy landmine concern and a looming threat of future contamination, as variants of these mines and their deployment systems remain in active service. The PFM-1 is composed of polyethylene plastic that presents a particularly difficult challenge to HMA operations.^{15,16} Other plastic anti-personnel mines have historically been composed of Type IV plastic, differing from the flexible polyethylene used for the PFM-1,¹⁵ which was notoriously widely used during the Soviet-Afghan War (1979–1989). Although some of these mines have deteriorated over time, many of the nearly ten million mines remain an active threat today due to cold climate preservation.¹⁷ These anti-personnel mines are designed to be ballistically dispersed from aluminum KSF-1S cartridges that contain four dispenser racks of eighteen mines with seventy-two mines in total (Figure 1). Guided by their stabilizing wings, the mines fall gently to the surface where they remain scattered in ellipsoidal minefields of 8–10 m x 18–20 m. While the original PFM-1 design was subsequently updated to include a self-destruction timer, modernized PFM-1S type mines fall short of the self-destruction criteria in Protocol II of the *Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons*,¹⁸ as studies show that nearly half of PFM-1S mines fail to self-detonate upon deployment.¹⁹ Many of the original PFM-1 stockpiles were destroyed by signatories of the APMBC, but considerable stockpiles of variants of the PFM-1 mines are thought to remain in arsenals of multiple countries.²⁰ Today, Russia hosts the world's largest stockpile of anti-personnel mines with an estimated 26.5 million.²¹ Moreover, as recently as 2019, the Russian army introduced the UMZ-G multipurpose tracked vehicle with minelaying capabilities compatible with PFM-1 bearing cassettes capable of dispersing nearly 20,000 PFM-1 type mines per hour.²²

Previous drone flights collected over inert PFM-1 mines in proxy environments in New York provided a critical proof of concept on the use of drones to aerially identify small plastic PFM-1 anti-personnel mines from visual, thermal, and multispectral imagery. Baur et al. 2020 presented a case study where a machine learning workflow was developed to automate the detection of these landmines for the first

time, resulting in a 71.5% accuracy rate in identifying and locating PFM-1 mines from a drone survey.²³ This study greatly improves upon the previous model by increasing the detection accuracy by over 20%. This was accomplished by increasing the volume of training data more than three-fold, now including twenty-five 10 x 20 m minefield orthophotos in four environments (rubble, sand, and two types of grass settings) with a total of 590 PFM-1 mines and 136 KSF casings photographed, adding negative samples, and a more customizable method of splitting orthomosaics.

In this study, we demonstrate that our convolutional neural network (CNN)-based results can be improved with the addition of new information from ground verification surveys or the addition of more imagery data. We present a specialized tool for detection of the PFM-1 anti-personnel landmine from RGB (visual) imagery, while providing

proof of concept and laying the foundation for other researchers to develop similar deep learning models for detection of other types of landmines, UXO, and ERW. In order to expand the capabilities of this particular CNN for automated detection of landmines or ERW other than the PFM-1, additional datasets of novel minefields would be required. With these data, our CNN has the potential to be generalized for automated detection and coordinate return of small objects in any raster dataset, including geophysical, RGB, thermal, and multispectral orthomosaics. Because of the self-learning capability of the neural network approach to detection and classification, it is important that researchers and field operators work together to develop methods for data sharing of aerial images, so the HMA community can continue to collectively refine and improve the reliability of deep learning methods that are rapidly gaining traction.²⁴

Background

Object Detection. The branch of machine learning used in this paper—object detection—involves the general goal of detecting and locating predetermined objects in often dynamic and complex environments. We used supervised object detection, which involves training an algorithm (a CNN in our case) to recognize objects by showing it a large number of images containing those objects (landmines in our case). The algorithm uses this set of images, the *training set*, to learn the distinguishing characteristics of the objects and form a generalized model in order to detect the objects in different images in the future. The algorithm is trained in *epochs* (rounds), where every image in the training set is analyzed and where the algorithm is tuned to be able to recognize images like it in the future. After the training of the algorithm is complete, a *pre-trained model* is created with the generalized model of the objects encoded in it. The accuracy of this model is evaluated using the *testing set*, containing images that were completely withheld from the training set. The lack of overlap in the training and testing sets is designed to ensure that the testing set provides an unbiased evaluation of the model, as it will not be trained to detect the objects as they appear in the testing set; it will have to rely on the generalized model of the objects it obtained from the training set.

The model is evaluated using three relative accuracy scores: precision (positive predictive value), recall (sensitivity), and F1 (harmonic mean). These scores are calculated from our raw evaluation metrics: true positives, false positives, and false negatives. Precision is a measure of the relative amount of positively labeled objects that were correctly labeled and is calculated using the following formula:

$$\text{Formula 1: } \text{prec} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}$$

Recall is a measure of the relative number of desired objects that were positively labeled and is calculated using the following formula:

$$\text{Formula 2: } \text{rec} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}$$

The F1 score is the harmonic mean of precision and recall; it is calculated using the following formula:²⁵

$$\text{Formula 3: } F1 = 2 * \frac{\text{prec} * \text{rec}}{\text{prec} + \text{rec}}$$

Algorithm 1. Coordinate computation

```

1: size = 700 (size of each cropped image)
2: overlap = 70 (overlap of cropped images in pixels)
3: cropped_image_name = 'Grass_RGB_Split1205.tif'
4: split_x, split_y = 370, 360 (pxl position of obj in cropped image)
5: ortho_easting, ortho_northing = 420000, 4600000 (UTM coords of ortho)
6: x_res, y_res = 0.00644, -0.00644 (m/pxl of orthophoto)
7:
8: (Get col and row of cropped image in ortho)
9: row = two ints following "Split" in cropped_image_name = 12
10: col = two ints following row in cropped_image_name = 5
11:
12: (Get pxl position of obj in respective orthomosaic)
13: ortho_x = split_x + (col * (size-overlap))
14: ortho_y = split_y + (row * (size-overlap))
15:
16: (Get easting and northing coords of obj)
17: obj_easting = ortho_easting + (ortho_x * x_res)
18: obj_northing = ortho_northing + (ortho_y * y_res)
19:
20: (Get latitude and longitude coords of obj - python utm package)
21: obj_lat, obj_lon = utm.to_latlon(obj_easting, obj_northing)

```

Algorithm 1. Pseudocode outlining coordinate computation of object.

For our implementation of object detection, we used an actively-maintained, open-source GitHub repository called Faster region-based convolutional neural network (R-CNN).²⁶ An important component of the Faster R-CNN architecture, which we modified to achieve a higher accuracy with our dataset, is the anchor box. These make it possible for the network to locate objects of different scales and different aspect ratios quickly and effectively. Instead of scanning an entire image using a sliding window of a set size and aspect ratio, nine anchor boxes are centered around every sixteenth pixel in the image. The default anchor boxes consist of all nine combinations of boxes with a ½:1, 1:1, and 2:2 aspect ratio, and boxes with a scale of 8, 16, and 32 multiplied by the default base anchor: 16.²⁷

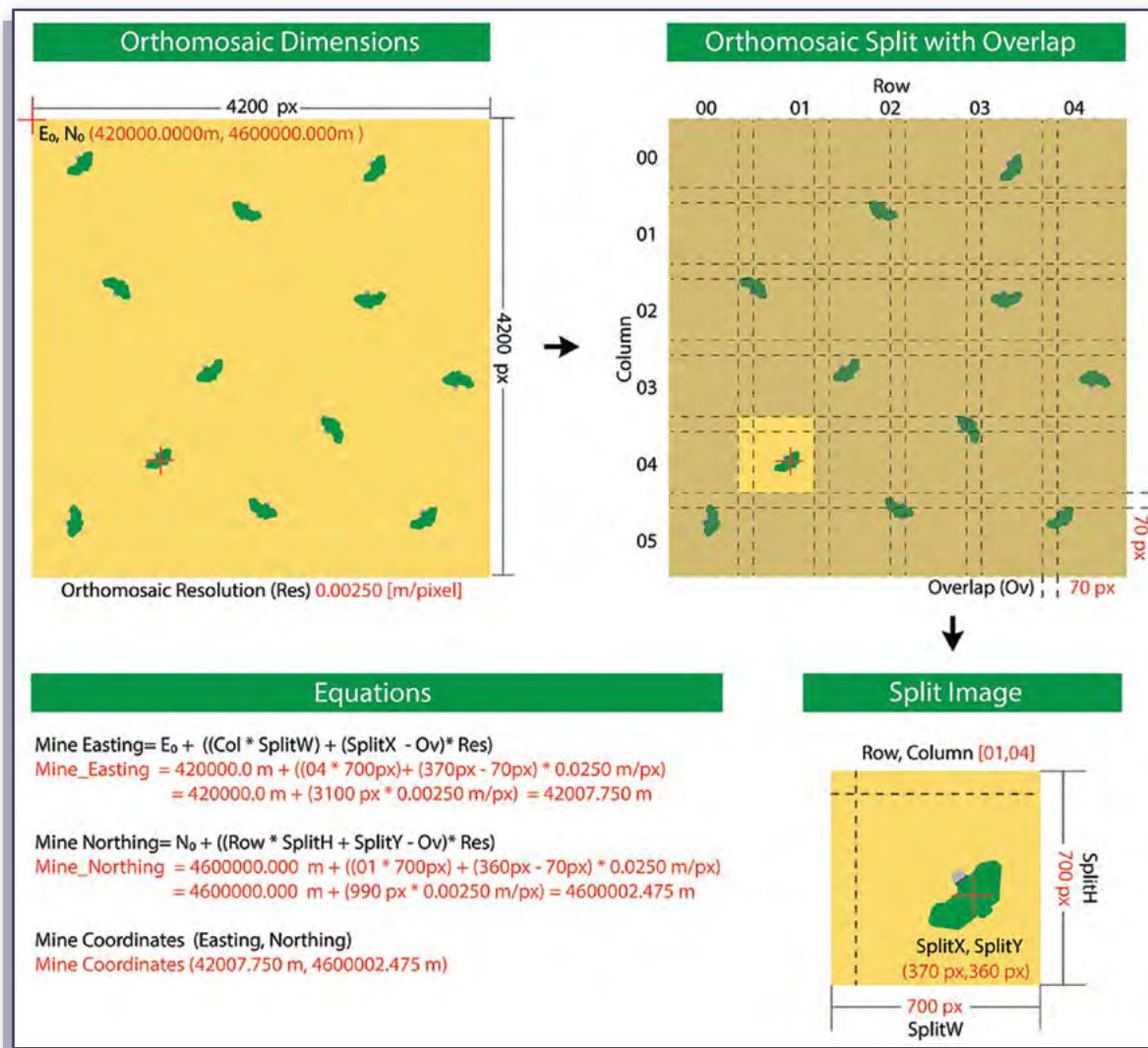


Figure 2. Illustration of how the CNN assigns coordinates to suspect mines.

Methodology

Remote sensing methodology. Additional training and testing data were acquired on the Binghamton University campus of an inert grass minefield and an inert sand training minefield. In both instances, the training minefields consisted of fifteen to thirty PFM-1 landmines and their KSF casings scattered throughout a 10 x 20 m region. We collected aerial images of the minefield with the DJI Phantom 4 Pro drone at a 10 m height and 2 m/s speed with an 80% overlap. The flights were planned using the Pix4D capture mission planner. In the grassy terrain, the mines were randomly dispersed so that their orientation was left to chance. In the sandy terrain, however, we specifically placed the mines in six different orientations to increase the robustness of our model: horizontal body up, horizontal body down, vertical cap up, vertical cap down, horizontal body buried (partial occlusion), and horizontal wing buried. We also collected control data over the sandy area with no mines to add negative samples (often called negative templates in remote sensing literature).²⁸ Additional orthomosaic simulated minefield data from rubble, grass, and snowy environments was used for this project and had previously been collected with the same

acquisition methods as presented here.²³ The addition of more data and negative samples greatly improves the accuracy of machine learning algorithms, critically decreasing false positives.

CNN improvements and adjustments. Since the previous implementation of a Faster R-CNN to detect PFM-1 landmines, our methods have changed and improved dramatically. Baur et al. 2020 describes the use of Impy to create non-overlapping crops of our orthomosaics with at least one object in each photo and corresponding XML files containing the bounding boxes around each object.²⁹ As in Baur et al. 2020, the Faster R-CNN begins the processing by resizing the inputted images to a maximum of 700 px for each axis. Therefore, we continue to employ the method of splitting orthomosaics but employ this method very differently, using a script we created called ImageSplitter.³⁰ This shift allows for the inclusion of negative samples (images with no objects) in our training and testing sets to improve our accuracy, provide more precise evaluation metrics, and most importantly, easily locate the predicted objects with latitude and longitude coordinates.

Impy, our previous splitting tool, split an image into non-overlapping 1032 x 1032 px crops each containing at least one bounding box. The two major limitations of this tool were the difficulty in labeling our predicted objects with latitude and longitude coordinates and the inability to create negative samples, which were not possible using Impy because of the requirement that there must be at least one object within each split image. It was difficult for us to locate our predicted landmines with coordinates using this tool, because the orthomosaics were split in a non-uniform way without the ability to output the offsets with which the split images relate to the orthomosaic as a whole. The location of the split images was created based on the location of randomly scattered landmines instead of on constant, predetermined offsets.

ImageSplitter solves both of these problems by splitting an image into square crops of a predetermined size (we chose 700 x 700 px, as larger images would be downsampled by the CNN) with a predetermined percent overlap (we chose 10%) and creating corresponding

XML annotations. This method allows us to input negative samples as training and testing data to boost our accuracy and to obtain a more precise evaluation of our model. The predetermined crop size and overlap also allows us to determine where each image crop would lie in the larger orthomosaic, and therefore allows us to locate the predicted objects with latitude and longitude as well as easting and northing coordinates. This process is outlined in Algorithm 1 and Figure 2.

After the calculations, the following data are written to a CSV (comma separated values) file, which can be read by Google Earth Pro or other geographic information systems (GIS) via the object name; the corresponding orthomosaic; its predicted confidence score (from 0 to 1; and its corresponding latitude, longitude, easting, and northing coordinates.

All training sessions were conducted over 50 epochs, with the Resnet-101 pre-trained model and an 8 px base anchor size with all other parameters set to the default values. The train-test splits that were used are outlined in the following results section.

Results

The experiments described in the methodology were done on a Dual Socket Intel® Xeon® Silver 4114 CPU at 2.20 GHz with 128 GB of RAM with a Titan V GPU with 12 GB of RAM. Thirteen experiments were executed by manipulating four variables.

We found that the highest mean F1 scores, regardless of train-test split, were achieved with a 700 x 700 px image size with a 10% overlap, 8 px base anchor size, and negative samples included. With these parameter values, we tested seven different train-test splits. These seven splits can be categorized in two ways: randomized splits and orthomosaic withholdings. The results from the experiments in each categorization give us unique insights. The randomized splits involved randomly selecting a percentage of images for the training set, while the withheld remainder are used for testing. This type of splitting gives us insight into how the model will perform when generating predictions on environments that have been partially included in the training set in the form of simulated minefields or negative samples. The

orthomosaic withholdings involve selecting an entire session of drone flights to make up the testing set, while the remainder are used for training. This train-test split gives us more accurate insight into how the model will perform when generating predictions on an environment that was completely excluded from the training set, but similar environments were present. The best randomized split obtained was an 80/20 train-test split (80% training, 20% testing) yielding a 0.95 F1 score for PFM-1 landmines, a 0.89 F1 score for KSF casings, and a 0.92 mean F1 score. This split is often standard, striking a balance between obtaining a robust model through training and a complete evaluation of the performance of the model during testing. The best orthomosaic withholding was obtained by withholding three orthomosaics we collected from a simulated rubble minefield in Chenango Valley State Park, NY. This split yielded a 0.86 F1 score for PFM-1 landmines, a 0.98 F1 score for KSF casings, and a 0.92 mean F1 score.

Using the Pre-Trained Model for PFM-1 Detection

Materials. Materials needed to detect and map PFM-1 anti-personnel mines using this method include a commercial off-the-shelf (COTS) drone, an RGB visual camera mounted on the drone, orthomosaic generation software, and GIS software to analyze and visually inspect drone imagery.

Steps. The first step to implementing our pre-trained model for automated detection of PFM-1 anti-personnel mines is collecting UAS flight data. Optimal data acquisition involves flying a UAS at a 10 m height (the chosen height for optimal resolution and ground coverage) with 80% overlap coverage between consecutive traverses and at

Materials needed to detect and map PFM-1 anti-personnel mines using this method include a commercial off-the-shelf (COTS) drone, an RGB visual camera mounted on the drone, orthomosaic generation software, and GIS software to analyze and visually inspect drone imagery.

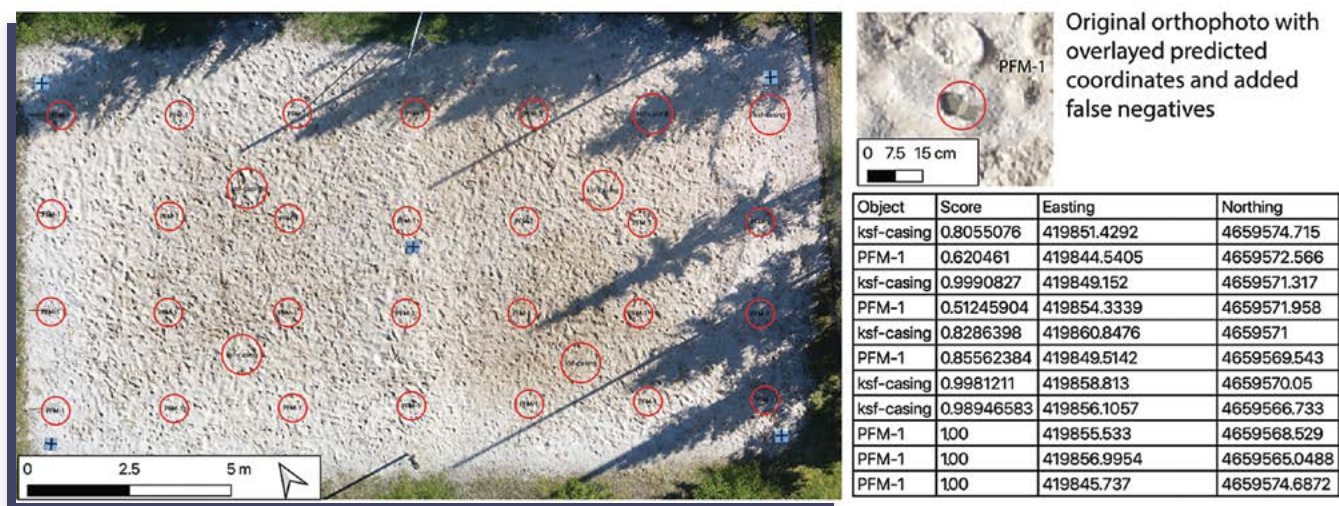


Figure 3. Edited shapefile overlaid on original orthomosaic with corresponding CSV file.

a slow speed (we used 2 m/s) to minimize blur in the images. Our survey size was 10 x 20 m in correspondence with the approximate dimensions of PFM-1 minefields, but this variable is up to the discretion of the operator and is often region specific. As the survey size increases, the processing time for constructing orthomosaics will also increase. The UAV must be equipped with an RGB-capturing camera such as the built-in camera on the DJI Phantom 4 or an external sensor such as the Parrot Sequoia multispectral sensor. Additional wavelengths, such as thermal or multispectral bands, may also be collected and used to cross-reference with visual imagery providing a multiparameter approach to confirm areas of potential surface-laid mines.²³ A multiparameter sensor system would help reduce false positives and provide additional physical information (such as temperature anomalies for thermal infrared imaging) of suspect mines, adding to the robustness of the technique. At this point, while these additional wavelengths are useful for cross-referencing with visual imagery, they are not suitable for input into the CNN as it is only trained on RGB images and is not currently able to recognize PFM-1 landmines in other types of imagery.

After data acquisition, the collected photos must be uploaded into photogrammetry software for orthomosaic generation. We used Pix4DMapper,³¹ but other software such as DroneDeploy, Agisoft Photoscan, and ESRI Drone2Map for ArcGIS are also capable of this task. For our trials, the Phantom 4 Pro camera at a 10 m height produced a resolution of 0.27 cm/px in the orthomosaic with the internal drone GPS with the use of ground control points to further improve location accuracy.

The resulting orthomosaic will produce a TIFF file and a corresponding TFW file containing important locational metadata. This metadata must include the x and y meters or centimeters per pixel resolutions (ours yielded 0.27 cm/px) and the easting and northing value of the top-left pixel of the orthomosaic. After the orthomosaics have been created, they must be split in order to be input to the CNN to generate PFM-1 and KSF-casing coordinate predictions. This will be accomplished using the ImageSplitter tool. Once a directory with the split orthomosaic has been created, follow the directions in the Faster R-CNN repository to execute predictions and output CSV files with the predicted coordinates in latitude and longitude, and UTM (Universal Transverse Mercator) formats. Notably, the zone for the outputted easting and northing coordinates will be the same zone present in the TFW files corresponding to the orthomosaics.

After the CNN outputs the coordinates of suspect mines, it is recommended to convert the CSV file to a shapefile (SHP) and overlay the SHP onto the original minefield orthomosaic. This can be accomplished using most GIS software and can be done for free using open-source software like QGIS or Google Earth Pro. An added benefit is that these maps may be downloaded for offline use in the field. Next, the operator can reduce the number of false negatives and false positives by visually inspecting the orthomosaic with the overlaid coordinate predictions to confirm or reject the location of the suspect mines, and to add any additional unidentified mines (Figure 3). More detailed directions on editing the CSV with predicted mines can be found at the Demining Research Community's website (de-mine.com) under the Open Source->Instructions tab.

While this methodology has a 91.8% accuracy for visible PFM-1 mines in sand, grass, and rubble proxy environments, it is important to note the limitations in real world situations that would complicate detection using the CNN and RGB imagery.

How will pairing drones and this CNN improve demining practices?

Safety. Detecting PFM-1 mines remotely from drones can reduce the uncertainty associated with finding previously unidentified minefields on the ground. Additionally, orthomosaics with identified suspect PFM-1s (Figure 3) will allow deminers to visualize where these mines may be before stepping onto the minefield and provide them with better situational awareness.³⁵ This method acts as an NTS for area reduction to assess SHAs or CHAs quickly and safely before investing more time and resources with a technical survey.

Cost. One of the major advantages of this method is its ability to reduce costs associated with mine detection by constraining the SHAs. The software required for this method includes our open-source CNN (freely available at de-mine.com), GIS software for making maps (QGIS software and Google Earth Pro software, both free), and orthomosaic generation software. For the orthomosaic generation software applications, we used Pix4D costing US\$4,990 for a lifetime license, or \$3,500 per year with free trial available, but other software applications available on the market include ArcGIS Drone2Map, costing \$1,500 per year (requires ArcGIS). Additionally, open-source (free) orthomosaic generation software options exist including Meshlab, MicMac, and VisualSFM. The hardware required includes a COTS drone with an RGB camera (we used the DJI Phantom 4 Pro costing ~\$1,600) and a commercial laptop on which to run predictions. We used a Lenovo Yoga C740 with a 10th Generation Intel® Core i5-10210U Processor and 8 GB of RAM, which is sufficient to execute predictions, costing ~\$850. The lowest estimated cost for the materials of this method is \$2,000 to the highest estimated cost of \$20,000. The lowest cost assumes using a COTS drone with a built-in camera, relying on open-source freeware or free software trials, and excluding the cost of purchasing a sufficiently powerful computer on which to run the software. The highest cost assumes the operator is buying a professional specialized drone and camera while purchasing a lifetime license for a photogrammetry software application.

Time. Collection of aerial drone footage takes approximately **3.5 minutes** for a **200 m²** minefield, covering roughly **1,143 m²** in **20 minutes** before the battery needs to be replaced or recharged for a typical DJI Phantom 3, assuming the drone is flying at a 10 m height, at 2 m/s, with 80% overlap.¹⁶ Orthomosaic generation takes about **1 hour and 17 minutes** for a **200 m²** area on a 2.7 GHz Dual-Core Intel® Core i5 processor with 8 GB of RAM. Generating predictions using the CNN takes about **0.04 seconds per cropped image** on a Dual Socket Intel® Xeon® Silver 4114 CPU at 2.20 GHz with 128 GB of RAM with a Titan V GPU with 12 GB of RAM. Generating predictions takes about **6 seconds per cropped image** using the Lenovo® Yoga C740 with a 10th Generation Intel® Core i5-10210U Processor and 8 GB of RAM. Manually verifying the CNN mine predictions and labeling any false negatives in QGIS takes approximately **10 minutes** for **30 objects** in a 10 x 20 m minefield. In total, the estimated time required for this methodology from start to producing field maps is **2 days**. The first day will be used for data acquisition and drone operation, while the second day will be used for orthophoto generation, CNN predictions, and creating field-ready maps.

Accessibility. Our method allows for remote assessment in regions that can be physically inaccessible to survey, but we suggest the operator retains line of sight of the UAS at all times. In rugged terrain with large elevation shifts, mission planning software is necessary to preprogram global navigation satellite system (GNSS)-guided autonomous missions where waypoints are used in navigation to maintain constant altitude above ground level (AGL); alternatively, a laser altimeter can be used to maintain constant altitude AGL. Our method will have limited success in highly-vegetated regions and in detecting mines that are occluded in the RGB wavelengths of light, including buried minefields.

Limitations

While this methodology has a 91.8% accuracy for visible PFM-1 mines in sand, grass, and rubble proxy environments, it is important to note the limitations in real world situations that would complicate detection using the CNN and RGB imagery. Firstly, PFM-1 mines that have been buried by earth surface processes or are completely covered in vegetation/obscured from an aerial view are not detectable with this algorithm or optical imagery of any kind. Detecting buried mines is outside the scope of this paper, but recent studies show other UAS-based techniques such as thermal sensing show promise for detecting

shallowly buried mines in dry environments.^{32,33} Furthermore, there may be some visible landmines that the CNN fails to detect, as our algorithm is not 100% accurate. Because of this, it is important for an operator to double check the orthomosaic for mines. Additionally, our algorithm was trained on non-deteriorated model PFM-1 mines, so this may introduce a bias in the CNN, and is an avenue for future work to include images of decade-old, in-situ PFM-1 mines. Many of the mines we trained on were exposed to the outdoors or previously buried, being caked with mud or sand as we would expect in nature.

Additionally, over time, as the PFM-1 mines become more deteriorated, the probability that these mines will have either already detonated due to reaching their cumulative triggering pressure or will have experienced a casing breach resulting in neutralization or disarmament increases.³⁴ However, some PFM-1 mines from the Soviet-Afghan War remain active and present a particularly difficult target to identify, since our algorithm is tailored to identify more recently-laid mines. In regard to detecting half-buried or half exposed mines, this CNN was provided minimal training data for these types of images and was only able to detect one of twenty-six half buried mines in the testing orthomosaic. While this number is low, it is also promising in that the CNN was able to successfully extrapolate from the fully exposed mines and is an avenue for future work that will dramatically improve as the training images of partially buried mines increases. Overall, the main limitations of this method are in detecting PFM-1 mines that are not visible on the surface, and lack of training data

from real world minefields which is logistically difficult to obtain and an avenue for future work.

Our methodology is intended to assist and augment current mine detection practices, not replace them. However, employing this methodology can improve the safety of operators; increase the efficiency, speed, and accuracy of detection; and reduce costs for conditions where this method is effective.

This method has potential applications for detecting PFM-1 anti-personnel mines in confirmed hazardous areas (CHAs) and SHAs contributing to a NTS for area reduction and partial detection without deploying personnel on the ground. Using our method in this fashion can be extremely helpful, as even detecting a single landmine in a region will provide critical information for mapping areas of contamination and help decision makers prioritize areas based on the contamination density.


Bigger Picture

Our methodology is intended to assist and augment current mine detection practices, not replace them. However, employing this methodology can improve the safety of operators, increase the efficiency, speed, and accuracy of detection, and reduce costs for conditions where this method is effective.

This method has potential applications for detecting PFM-1 anti-personnel mines in both confirmed hazardous areas (CHAs) and SHAs contributing to a NTS for area reduction and partial detection without deploying personnel on the ground. Using our method in this fashion can be extremely helpful as even detecting a single landmine in a region will provide critical information for mapping areas of contamination and help decision makers prioritize areas based on the contamination density.

Conclusion

Recent advances in machine learning, miniaturization of sensors, and the commercialization of drones are paving the way for the future of automated mine detection. This study couples these powerful technologies by training a CNN on UAV-based minefield data, producing a model that can identify the PFM-1 anti-personnel mine from a drone survey with 91.8% accuracy and can provide deminers field-ready maps with identified mine locations. By following the steps outlined in this paper, deminers can successfully implement this CNN to automate detection of PFM-1 anti-personnel mines. This CNN can be adapted to automate detection of a range of landmines, cluster munitions, and other ERW, given adequate training and testing data. Our future work will involve field testing this method in active minefields to better understand the environmental parameters that may influence this methodology. While this methodology is not meant to replace current demining practices, it is capable of augmenting these practices by providing a safe, low-cost, time-efficient, and accurate detection method to add to the demining toolbox for CHAs and SHAs.

Data availability statement. The Demining Research Community is a group of interdisciplinary scientists with backgrounds in remote sensing, geophysics, computer science, and archaeology whose mission is to research, develop, and field test cutting-edge sensors and platforms to improve current demining technologies. We support open-access research and data for the betterment of the mine action community. Our minefield datasets (around 160 MB each), source code, and previous publications on mine detection are available at the Demining Research Community's website and at the Open Repository at Binghamton University.³⁶⁻³⁸ 

Acknowledgments. Our research team would like to thank the First-year Research Immersion and Harpur Edge for their support of the project. This work was conducted under New York State Parks Unmanned Aircraft and Special Use permits, and we extend our gratitude to park manager Michael Boyle and all staff of the Chenango Valley State Park for their assistance with this project. We would also like to thank the reviewers and editors of this paper for their insightful comments and suggestions that strengthened the manuscript greatly.

See endnotes page 154

Jasper Baur

Columbia University - Lamont-Doherty Earth Observatory



Jasper Baur is a Ph.D. student in the Department of Earth and Environmental Sciences at Columbia University studying volcanology with a specialization in remote sensing. He is the founder of the Demining Research Community, which researches, develops, and field tests sensors and platforms to improve demining technologies. Baur received his B.S. in geological sciences at Binghamton University, and double minored in graphic design and geographic information systems.

Gabriel Steinberg

Binghamton University - The State University of New York



Gabriel Steinberg completed his Computer Science B.S. at Binghamton University and is preparing to begin his M.S. in Computer Science. Steinberg's main research interest is in object detection with the goal of detecting PFM-1 landmines in post-conflict nations, especially Afghanistan, and he has published that research in *Remote Sensing* as second author. Steinberg has worked as an Undergraduate Research Assistant at UC San Diego and at the Hamburg University of Applied Sciences.

Alex Nikulin, Ph.D.

Binghamton University - The State University of New York



Alex Nikulin, Ph.D., is an Associate Professor and Director of the Humanitarian Geophysics Research Program with the Department of Geological Sciences and Environmental Studies. Nikulin is a co-founder of the Geophysics and Remote Sensing Research Laboratory at Binghamton University, a collaborative academic platform that brings together undergraduate and graduate researchers, and faculty, as well as industry, academic, government, and NGO partners, to work on resolving issues facing society through applications of geophysical techniques.

Kenneth Chiu, Ph.D.

Binghamton University - The State University of New York



Kenneth Chiu, Ph.D., is an Associate Professor in the Department of Computer Science at Binghamton University. His research interests are in the area of high-performance computing, applied machine learning, and cyberinfrastructure. His work has been funded by agencies including, DOE, NIH, and NSF. He holds degrees in Computer Science from Princeton University and Indiana University.

Timothy S. de Smet, Ph.D.

Binghamton University - The State University of New York



Timothy S. de Smet, Ph.D., is a Research Assistant Professor in the Department of Geological Sciences and Environmental Studies, Director of the Geophysics and Remote Sensing Laboratory, and leader of the Environmental Visualization with Drones research stream in the First-year Research Immersion at Binghamton University. His areas of expertise are aerial remote sensing and near-surface applied geophysics. Dr. de Smet's research utilizing frequency and time-domain electromagnetic-induction, magnetometry, ground-penetrating radar, aerial LiDAR, and thermal infrared remote sensing has been published in *Geophysics*, *Remote Sensing*, *Near Surface Geophysics*, *The Journal of Applied Geophysics*, *Sedimentary Geology*, *PLOS ONE*, *The Journal of Conventional Weapons Destruction*, *Archaeological Prospection*, and *The Leading Edge*, among others. Dr. de Smet is an FAA 107 certified UAS remote pilot.

Watch out for

The CISR EXCHANGE

Conversations with contributors to
The Journal of Conventional Weapons Destruction,
and experts in humanitarian mine action

Learn more by emailing cisr-journal@jmu.edu

ENDNOTES

How to Implement Drones and Machine Learning to Reduce Time, Costs, and Dangers Associated with Landmine Detection
By Jasper Baur,i,ii* Gabriel Steinberg,i* Alex Nikulin, Ph.D.,i Kenneth Chiu, Ph.D.,i and Timothy S. de Smet, Ph.D.i

i[Binghamton University - The State University of New York]

ii[Columbia University - Lamont-Doherty Earth Observatory]

*Equal Contribution

1. "Deep learning is a subfield of machine learning based on the use of deep neural networks. Machine learning is a subfield of artificial intelligence wherein a computer uses algorithms to improve at a task on its own only through experience. Artificial intelligence describes any program designed to imitate human perception.
2. A region-based convolutional neural network (R-CNN) is a method introduced in 2014 to perform fast and accurate object detection using deep neural networks.
3. International Campaign to Ban Landmines. (2020). *22nd annual Landmine Monitor Report*. Human Rights Watch.
4. Mine Action Review. (2020, October 1). CLEARING THE MINES 2020. Retrieved from https://reliefweb.int/sites/reliefweb.int/files/resources/907_NPA_Clearing_the_Mines_2020_WEB.pdf
5. Habib, M. K. (2007). Humanitarian demining: reality and the challenge of technology—the state of the arts. *International Journal of Advanced Robotic Systems*, 4(2), 19.
6. *Landmine Monitor 2020*, (2020), p 1, <http://www.the-monitor.org/media/3168934/LM2020.pdf>.
7. UAV refers solely to the aerial platform (drone), whereas UAS refers to the drone along with the attached sensors, the ground control station, and the remote pilot controlling the autonomous flight.
8. de Smet, T., & Nikulin, A., (2018) Catching "butterflies" in the morning: A new methodology for rapid detection of aerially deployed plastic land mines from UAVs. *The Leading Edge*, 37(5), 367-371.
9. deSmet, T., Nikulin, A., Frazer, W., Baur, J., Abramowitz, et al. (2018). Drones and" Butterflies": A Low-Cost UAV System for Rapid Detection and Identification of Unconventional Minefields. *Journal of Conventional Weapons Destruction*, 22(3), 10.
10. Nikulin, A., De Smet, T. S., Baur, J., Frazer, W. D., & Abramowitz, J. C. (2018). Detection and identification of remnant PFM-1 'Butterfly Mines' with a UAV-based thermal-imaging protocol. *Remote Sensing*, 10(11), 1672.
11. Nikulin, A., de Smet, T., Puliaiev, A., et al. (2020). AUTOMATED UAS AEROMAGNETIC SURVEYS TO DETECT MBRL UNEXPLODED ORDNANCE. *Journal of Conventional Weapons Destruction*, 24(1), 13.
12. Fardoulis, J. S., Smith, D. A., Payton, O. D., Scott, T. B., Freer, J. E., & Day, J. C. C. (2017). Deploying Low Cost, Small Unmanned Aerial Systems in Humanitarian Mine Action. In *Proceedings of the 14th International Symposium "MINE ACTION"*.
13. EMI occurs when a current applied to a coil induces a magnetic field. This magnetic field will then induce electrical eddy currents when exposed to a metallic object, which then results in the induction of its own magnetic field towards the coils, generating an opposite current in the coil. This current from the metallic object is then detected by the instrument.
14. Nikulin, A., & de Smet, T., (2018). A UAV-based magnetic survey method to detect and identify orphaned oil and gas wells. *The Leading Edge* 38 (6), 447-452.
15. Harmon, R. S., DeLucia, F. C., LaPointe, A., Winkel, R. J., & Miziolek, A. W. (2006). LIBS for landmine detection and discrimination. *Analytical and bioanalytical chemistry*, 385(6), 1140-1148.
16. Coath, J.A.; Richardson, M.A. (2000). Regions of high contrast for the detection of scatterable land mines. *Proceedings of the Detection and Remediation Technologies for Mines and Mine-like Targets V*, Orlando, FL, USA, 24–28. Volume 4038, pp. 232–240.
17. Urban, M. (1990). Nine: 1986. In *War in Afghanistan* (pp. 186-211). Palgrave Macmillan, London.
18. United Nations. (1998, December 3). *Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices*. Available online: <https://treaties.un.org/doc/Publication/MTDSG/Volume%20II/Chapter%20XXVI/XXVI-2-b.en.pdf> (accessed on October 10, 2020).
19. Dolgov, R. (2001). Landmines in Russia and the former Soviet Union: A lethal epidemic. *Medicine & Global Survival*, 7(1), 38-42.
20. "AP Mine Ban Convention: Destroying Stockpiled Mines". (2017) [Apminebanconvention.org](https://www.apminebanconvention.org) (accessed 10 October 2017). <https://web.archive.org/web/20170826034242/https://www.apminebanconvention.org/background-status-of-the-convention/destroying-stockpiled-mines/>.
21. *Landmine Monitor*. (2019). International Campaign to Ban Landmines - Cluster Munition Coalition. <http://www.the-monitor.org/media/3074086/Landmine-Monitor-2019-Report-Final.pdf>
22. Army Recognition. Army-2019: New UMZ-G Multipurpose Tracked Minelayer Vehicle Based on Tank Chassis. Available online: https://www.armyrecognition.com/army-2019_news_russia_online_show_daily_media_partner/army-2019_new_umz-g_multipurpose_tracked_minelayer_vehicle_based_on_tank_chassis.html (accessed on 15 January 2020).
23. Baur, J., Steinberg, G., Nikulin, A., Chiu, K., & de Smet, T. S. (2020). Applying Deep Learning to Automate UAV-Based Detection of Scatterable Landmines. *Remote Sensing*, 12(5), 859.
24. Jebens, M., Shen, J., & Tollefsen, E. (2020). To What Extent Could the Development of an Airborne Thermal Imaging Detection System Contribute to Enhance Detection?. *The Journal of Conventional Weapons Destruction*, 24(1), 14.
25. Kent, A., Berry, M., Luehrs, F., & Perry, J.W. (1955). Machine literature searching VIII. Operational criteria for designing information retrieval systems. *American Documentation*, 6(2), 93.
26. Github. GSteinberg/Faster-RCNN.PyTorch. Available online: <https://github.com/GSteinberg/faster-rcnn.pytorch> (accessed on 4 October 2020).
27. Ren, S., He, K., Girshick, R., Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.*, 39, 91-99.

28. Negative samples (or negative templates) are images collected with no objects of interest present (PFM-1 mines or KSF casings in this case). Including them in the training data can often decrease the number of false positives by allowing the computer to clearly distinguish between background and objects of interest.
29. Github. Lozuwa/Impy. Available online: <https://github.com/lozuwa/impypy#images-are-too-big> (accessed on 4 October 2020).
30. Github. GSteinberg/ImageSplitter. Available online: <https://github.com/GSteinberg/ImageSplitter> (accessed on 4 October 2020).
31. Pix4D. Professional photogrammetry and drone mapping software. Available online: <https://www.pix4d.com/> (accessed on 4 October 2020)
32. Fardoulis, John; Depreytere, Xavier; Gallien, Pierre; Djouhri, Kheria; Abdourhmane, Ba; and Sauvage, Emmanuel. (2020). Proof: How Small Drones Can Find Buried Landmines in the Desert Using Airborne IR Thermography, *The Journal of Conventional Weapons Destruction*: Vol. 24 : Iss. 2 , Article 15.
33. Jebens, Martin; Sawada, Ph.D., Hideyuki; Shen, Junjie; and Tollefsen, Erik. (2020). To What Extent Could the Development of an Airborne Thermal Imaging Detection System Contribute to Enhance Detection? *The Journal of Conventional Weapons Destruction*: Vol. 24 : Iss. 1 , Article 14.
34. Ressler, Daniele, "Study of the Effects of Aging on Landmines" (2010), 1.
35. Fardoulis, J., Depreytere, X., Sauvage, E., & Gallien, P. (2019). Drones in the Desert: Augmenting HMA and Socio-Economic Activities in Chad. *The Journal of Conventional Weapons Destruction*, 23(1), 16.
36. de Smet, Timothy; Nikulin, Alex; and Baur, Jasper, "Scatterable Landmine Detection Project Dataset 1-7" (2020). *Geological Sciences and Environmental Studies Faculty Scholarship*. 1-7. https://orb.binghamton.edu/geology_fac/4
37. Steinberg, Gabriel; Baur, Jasper; Nikulin, Alex; and de Smet, Timothy, "Scatterable Landmine Detection Project Dataset 8-9" (2020). *Geological Sciences and Environmental Studies Faculty Scholarship*. 8-9. https://orb.binghamton.edu/geology_fac/11
38. Baur, Jasper; de Smet, Timothy; Nikulin, Alex; and Steinberg, Gabriel, "Scatterable Landmine Detection Project Datasets 10 - 24" (2020). *Geological Sciences and Environmental Studies Faculty Scholarship*. 10 - 24. https://orb.binghamton.edu/geology_fac.