SERDP

FINAL REPORT

Optical Detection and Classification of Military Munitions Underwater

Jed Wilbur
Jules Jaffee
Creare

June 2021

SERDP Project MR19-1423

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
21/06/2021	SERDP Final Report	7/22/2020 - 7/21/2021

4. TITLE AND SUBTITLE

Optical Detection and Classification of Military Munitions Underwater

5a. CONTRACT NUMBER
20-P-0064

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jed Wilbur
Jules Jaffee

5d. PROJECT NUMBER
MR19-1423

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Creare LLC
16 Great Hollow Road
Hanover, NH 03755

8. PERFORMING ORGANIZATION REPORT NUMBER
MR19-1423

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Strategic Environmental Research and Development Program (SERDP)
4800 Mark Center Drive, Suite 16F16
Alexandria, VA 22350-3605

10. SPONSOR/MONITOR'S ACRONYM(S)
SERDP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
MR19-1423

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Many active and former United States military installations are contaminated with munitions, including underwater sites. Of particular concern are shallow sites with exposed munitions. Optical imaging techniques offer promise for detecting these munitions: the imaging process is less affected by clutter; information such as optical contrast, color, and exact geometry is preserved, providing valuable inputs for automatic target recognition; imaging resolution is high enough to assess the integrity of the munition; and optical images are natural and intuitive for users to assess. Here we discuss the development of a proof-of-concept Optical Munition Detector (OMD) The OMD uses optical imaging techniques to survey the seafloor, extracting information useful for munition detection and classification.

The overall objective of this effort is to develop an Optical Munition Detector (OMD) that uses optical imaging techniques to detect and classify munitions underwater. The OMD is designed to close operational gaps with existing remedial investigation technologies. The specific objective of this limited scope project is to design, fabricate, and demonstrate a proof-of-concept OMD.

15. SUBJECT TERMS

Optical Detection, Military Munitions Underwater, UXO, UXO in the underwater environment, cluttered environment, target identification, UXO detection and classification, OMD< optical munition detector

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Jed Wilbur
UNCLASS	UNCLASS	UNCLASS	UNCLASS	64	**19b. TELEPHONE NUMBER** *(Include area code)* 603-643-3800

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

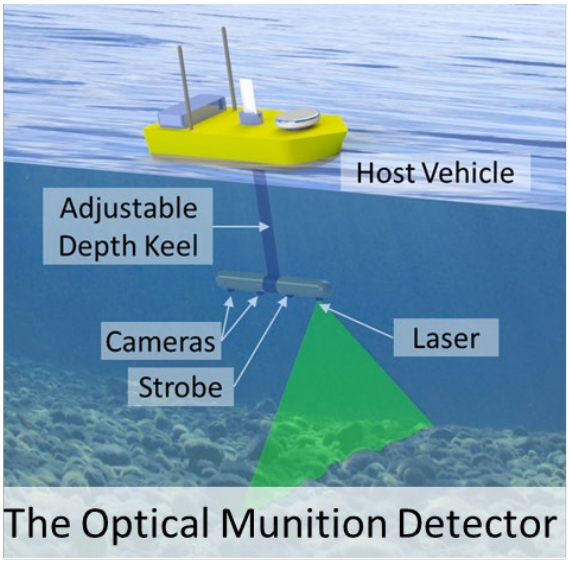
1 ABSTRACT

Introduction and Objectives. Many active and former United States military installations are contaminated with munitions, including underwater sites. Of particular concern are shallow sites with exposed munitions. Optical imaging techniques offer promise for detecting these munitions: the imaging process is less affected by clutter; information such as optical contrast, color, and exact geometry is preserved, providing valuable inputs for automatic target recognition; imaging resolution is high enough to assess the integrity of the munition; and optical images are natural and intuitive for users to assess. Here we discuss the development of a proof-of-concept Optical Munition Detector (OMD) The OMD uses optical imaging techniques to survey the seafloor, extracting information useful for munition detection and classification.

The overall objective of this effort is to develop an Optical Munition Detector (OMD) that uses optical imaging techniques to detect and classify munitions underwater. The OMD is designed to close operational gaps with existing remedial investigation technologies. The specific objective of this limited scope project is to design, fabricate, and demonstrate a proof-of-concept OMD.

Technical Approach. The OMD uses two mature and complementary optical metrology techniques: structured light and structure from motion. The first approach uses triangulation of a coherent light source and a camera to reconstruct the three-dimensional scene; the second extracts matched features in multiple views over a scene to determine relative camera positions so the three-dimensional scene can be reconstructed via stereo vision techniques.

Results. We demonstrated the feasibility of the OMD by designing and fabricating a proof-of-concept system and collecting and analyzing data of inert munitions in a local lake. Image quality was sufficient to construct 3D scenes from both optical modalities. We investigated and demonstrated a number of promising techniques for automatic munition detection from OMD data. In particular we developed a deep learning approach that does not require the extensive and expensive training typically necessary. We identified a number of technical areas with room for improvement in a next generation design. Finally, we investigated how the OMD may integrate with and improve existing munition remediation activities.



Benefits. The OMD is designed to close operational gaps in the remedial investigation of underwater sites suspected of munition contamination. Specifically, the OMD is designed to provide information on munition location, type, and integrity under conditions that are difficult for currently fielded technologies. This additional data will improve the performance of automatic target recognition and will improve the overall quality of data site managers have when making plans and decisions related to remedial action.

2 EXECUTIVE SUMMARY

2.1 INTRODUCTION

Many active and former United States military installations are contaminated with munitions, particularly unexploded ordnance (UXO) and discarded military munitions (DMM). The contaminated locations often include underwater sites such as ponds, lakes, and coastal ocean areas. Of particular concern are shallow sites with exposed munitions. To understand this hazard and to guide remedial action, improved survey tools are needed to assess the locations of munitions and their condition. Traditionally, underwater munitions are detected and classified using acoustic or electromagnetic techniques. However, there are gaps in the operational performance of these modalities: imaging is challenged by clutter; system outputs are a strong function of munition orientation (challenging classification algorithms); the results require a trained user to interpret; and little information is offered on munition condition (structural integrity). Optical imaging techniques offer promise for closing these gaps: the imaging process is less affected by clutter; information such as optical contrast, color, and exact geometry is preserved, providing valuable inputs for automatic target recognition; imaging resolution is high enough to assess the integrity of the munition; and optical images are natural and intuitive for users to assess, providing a welcome check on automatic target recognition hits. Here we discuss the development of a proof-of-concept Optical Munition Detector (OMD, Figure 1). The OMD uses optical imaging techniques to survey the seafloor, extracting information (high-resolution, three-dimensional topology, optical contrast, and color) useful for munition detection and classification.

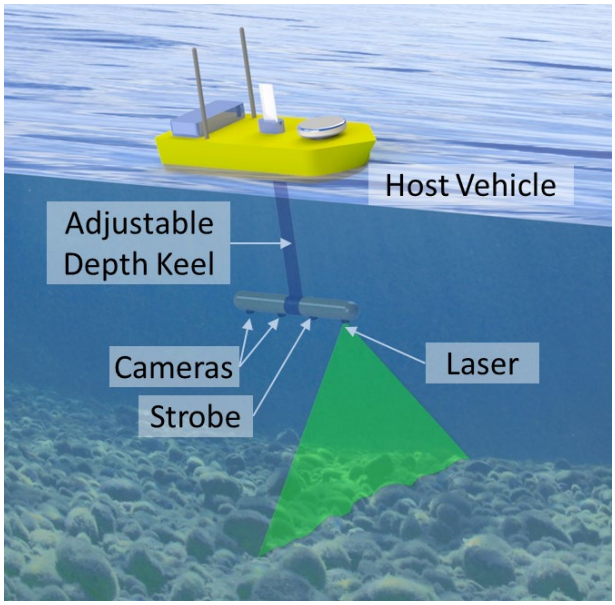


Figure 1. The Optical Munitions Detector

2.2 OBJECTIVES

The overall objective of this effort is to develop an Optical Munition Detector (OMD, Figure 1) that uses optical imaging techniques to detect and classify munitions underwater. The OMD is designed to close operational gaps with existing remedial investigation technologies. The

overall objective of this limited scope effort was to design, fabricate, and demonstrate a proof-of-concept OMD. The specific objectives of this effort were:

- Develop a set of specifications for an optical instrument to fill SERDP's operational gaps;
- Design and fabricate a proof-of-concept demonstration system;
- Demonstrate, in representative open water conditions, the ability to scan munition-mimicking targets using data collected by the proof-of-concept system;
- Research candidate algorithms for automatic target recognition (ATR) well suited for OMD data; and
- Develop the high-level design for a fully functional prototype.

2.3 TECHNICAL APPROACH

The Optical Munitions Detector uses optical imaging techniques to survey the bottom of the water column, extracting information (high-resolution, three-dimensional topology, optical contrast, and color) useful for munition detection and classification. The OMD has a number of advantages over traditional underwater acoustic imaging:

- It uses lower power (order tens of Watts), making it well suited for long deployments on unmanned vehicles;
- It has high resolution (sub millimeter), allowing for detailed characterization of munition condition and integrity;
- It is sensitive to optical contrast (e.g. markings and lettering) and color, features that will be very useful for ATR and human confirmation of positive ATR hits;
- It preserves the target object's geometric shape, another very useful feature for ATR and human confirmation of positive ATR hits;
- It provides immunity to multipath scattering, making it particularly well suited for shallow water missions;
- Its imaging process is less confounded by bottom clutter;
- OMD images are well suited to leverage the tremendous amount of ongoing effort of image classification via deep learning;
- Its output images are naturally and intuitively human-readable, allowing for a simple human-in-the-loop check for false positives on RI data before beginning RA;
- It is platform agnostic and can be deployed on surface vessels (manned or unmanned, autonomous or remote) or undersea vehicles.

OMD uses two mature and complementary optical metrology techniques: structured light and structure from motion. The first approach uses triangulation of a coherent light source and a camera to reconstruct the three-dimensional scene; the second extracts matched features in

multiple views over a scene to determine relative camera positions so the three-dimensional scene can be reconstructed via stereo vision techniques.

We divided the technical work on this project into four tasks: system specifications, proof-of-concept design and fabrication, proof-of-concept demonstration, and recommendations for a next-generation system.

2.4 Results and Discussion

2.4.1 System Design and Fabrication

System Design. We identified key performance parameters for the OMD optical components. In particular, resolution requirements drive camera pixel count and frame rate. Small rounds, on the order of 20mm, should be clearly resolvable in the data. Practically, we interpret this to mean their images should span at least 10-15 pixels. For similar reasons, the spatial offset between subsequent images should be on the order of 5-10mm, indicating that high speed imaging is necessary for reasonable area coverage rates. Reasonable underwater fields of view (to keep distortion low) can go up to roughly 60°. Based on this assumption, target per-pixel resolution, and the desire to maximize scan area, we targeted a scan altitude of approximately 1.5m and a swath width on the order of 1.5-2.0m.

System Fabrication. Figure 2 shows our assembled proof-of-concept system. This system was fabricated to our design with one notable exception. As we had to rush to stay ahead of our test site freezing over, the bandpass filter for the monochrome structured light camera did not arrive in time for incorporation into our test setup.

Figure 2. The Assembled Proof-of-Concept System (Left) and Replica Inert Munitions (Right)

2.4.2 System Demonstration

Munition-Mimicking Targets. As our system is sensitive only to shape, size, color, and texture, we purchased inert replicas from an online source (Inert Products LLC <inertproducts.com>). Figure 2 (Right) shows the inert munitions purchased in support of this effort:. A 105mm Artillery Projectile, two 25mm Bushmaster Rounds, an 81mm Mortar, a 155mm Artillery Projectile, an 81mm Mortar (with UXO texture and coloring), and a 105mm Artillery Projectile (with UXO texture and coloring).

Test Execution. We tested the proof-of-concept system at a local lake (Mascoma Lake, Enfield, NH) in November and December of 2020. The first (November) event was a learning exercise that greatly influenced the design and execution of the second (December) event. We obtained clear images from both the SfM and structure light subsystems (Figure 3). Only data from the second test event were analyzed.

Figure 3. Selected Color Images from the Second Test Event

2.4.3 Results – Lake Floor Topology

Structure from Motion. Figure 4 shows a detail of an example structure from motion (SfM) scene from one data set. This data set applied the camera calibration to produce results in engineering units (e.g. meters). The four munitions in the scene are clearly visible and properly shaped. Vegetation is also clearly visible in the scene.

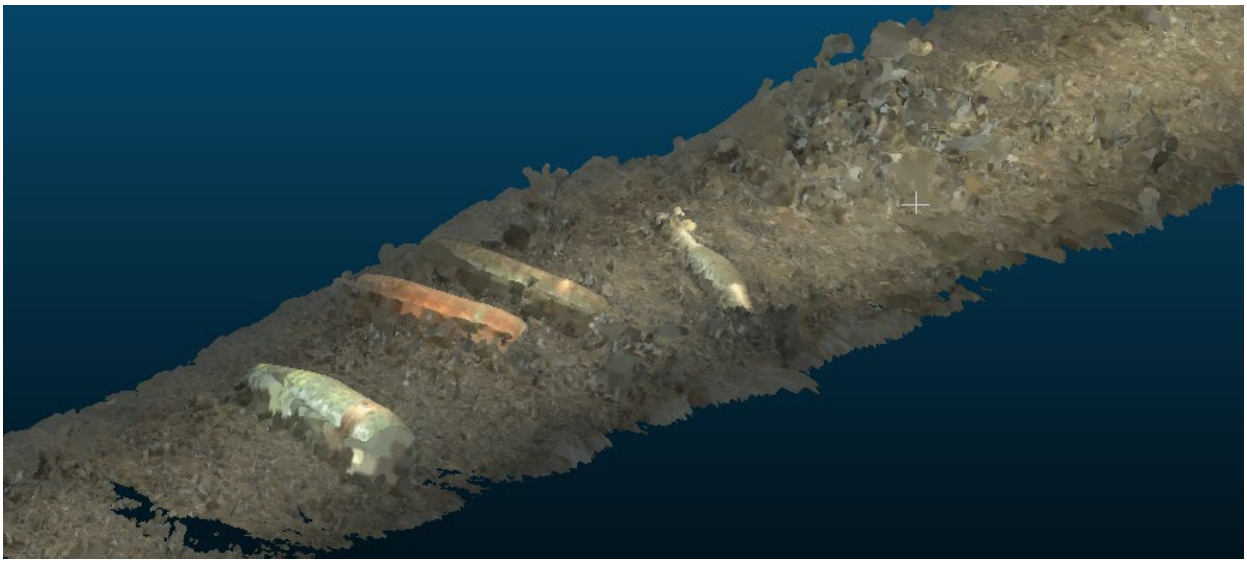


Figure 4. Detail of an Example Reconstructed Structure from Motion Scene from an Entire Imaging Run. Four Munitions are Clearly Visible.

Structured Light Imaging. We synthesized 3D measurements of our lake test scene from the structured light data. This effort uses laser-line segmentation in conjunction with the calibrated system model to construct a measurement through triangulation. Each image in the dataset gives a

line of three-dimensional points, which are then concatenated in a manner consistent with the estimated speed of our test vessel. The result from one pass over our test field is shown in Figure 5. The UXO targets present in the field of view are clearly visible in the data, and the structured light imaging approach results in a quantitative characterization of their size, shape, and orientation.

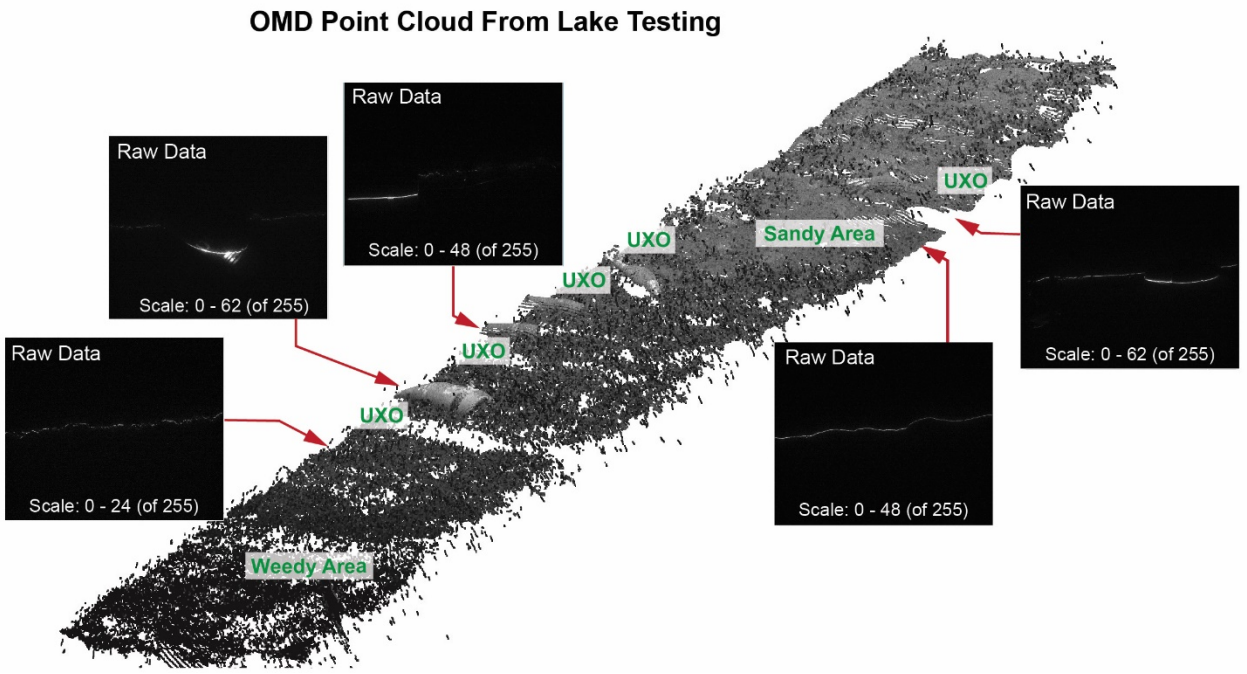


Figure 5. A Three-Dimensional Point Cloud Measurement of Our Test Scene.

2.4.4 Results – Improvements for the Next-Generation System

We identified a number of areas where performance could be improved. In particular, the lighting of the structure from motion scene can be improved with better lighting to: improve uniformity, reduce shadowing, improve power efficiency, and reduce backscatter.

2.4.5 Results –Target Detection Algorithms

We investigated a number of candidate algorithms to automatically detect munitions in OMD data. Figure 6 summarizes the results of this testing. Image texture analysis looks at the variation in image intensity computed over small regions of the image. While results are good for this data set, we are worried they will not be generalizable to the real-world. Curvature uniformity looks at regions with near constant curvature in three dimensions (using the SfM data set). This approach is promising as it can be tailored for munition specific shapes and should be fairly immune to natural clutter. We also looked at the probability that a deep convolution neural network (DCNN) trained (on another effort) to classify images as being “natural” or “man-made”. Despite being trained on in-air data with no munitions, this algorithm does a very good job of automatically detecting munitions; we present a number of approaches for improving algorithm performance.

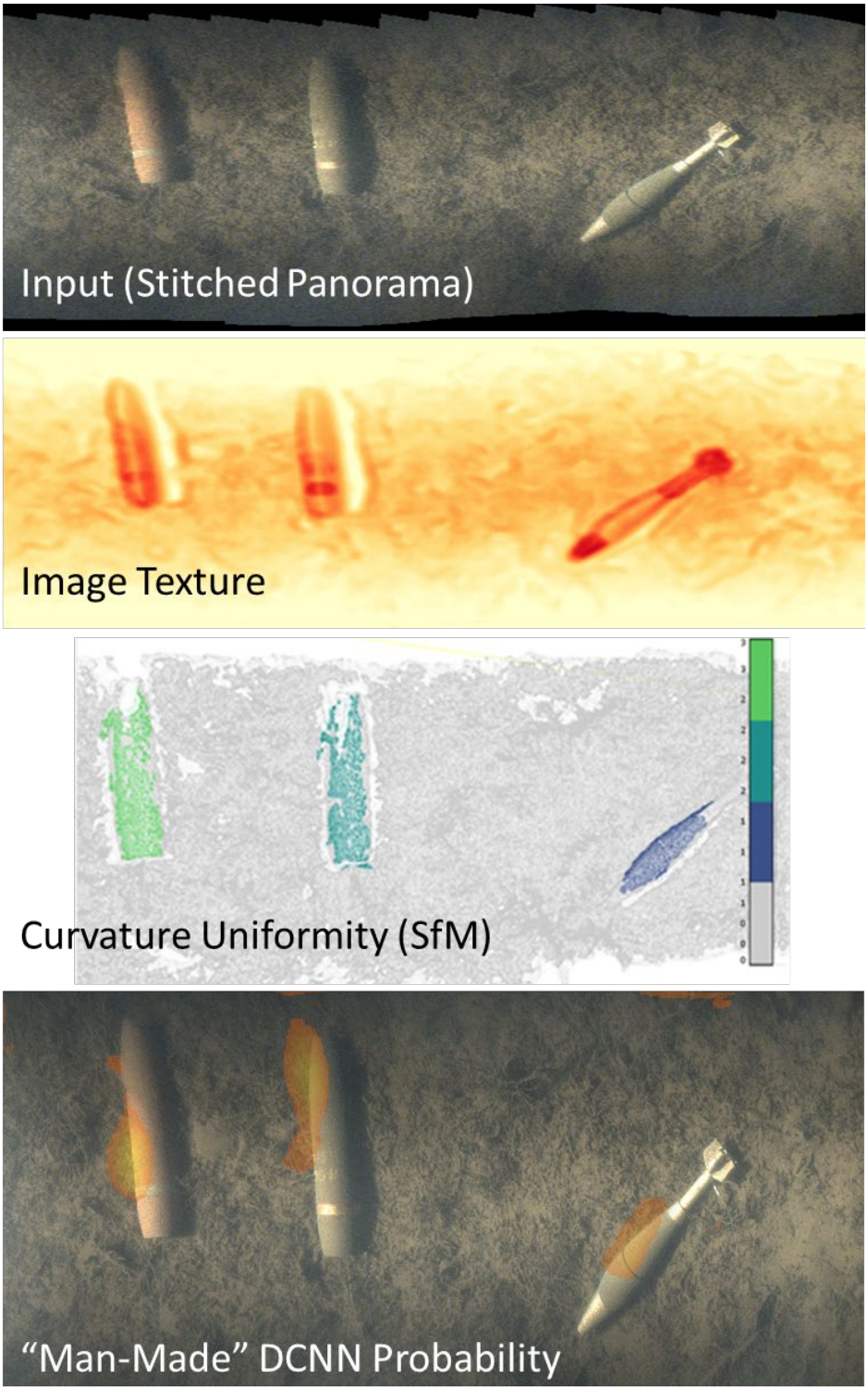


Figure 6. Initial Results for Sample Munition Detection Approaches. From top to bottom: raw camera scene; munition likelihood via image texture analysis; munition detection via curvature uniformity analysis of the SfM point clouds; munition likelihood via a deep convolution neural network trained to find “man-made” objects.

Finally (not show in the Figure), we developed (but have not yet tested) an approach to use DCNNs to detect munitions in structured light data. The approach avoids the need for manual labeling of training data by using the known geometries of the target munition. This approach is worthy of testing in a follow-on effort.

2.4.6 Conclusions

We drew a number of conclusions and lessons learned from this work:

- Our general imaging and reconstruction approaches are feasible and well suited for exposed munitions in shallow-water environments. Although the data set is, as anticipated, limited, munitions are clearly visible in the raw color images, composite structure-from-motion point clouds, and structured-light point clouds.

- The illumination of the structure-from-motion (color camera) scene needs improvement. There are strong lighting gradients, strong shadowing, and particle-backscatter that confound the SfM reconstruction and DCNN munition detection algorithms. Recommendations for improving the lighting (and image quality in general) should be implemented in the next-generation system.

- The signal to noise ratio of the structured light imaging system was, as expected given the hardware we used, poor. Increasing the laser power and the addition of a band-pass filter will vastly improve this. This has low technical risk.

- A number of munition detection algorithms have promise. We have identified shape-based approaches for the SfM point clouds; texture-based approaches for the raw images; machine learning (DCNNs) for the raw camera images; and DCNNs for the structured light data.

- Of these, the DCNN algorithms are likely the most generalizable. Training neural networks is typically the most time-consuming and expensive part of the development process. As we have shown here, applying networks trained on existing data sets (in-air, no munitions) can serve as a very reasonable starting point if the network outputs are appropriately categorized. These can be retrained on-the-fly with human-in-loop review of the data.

2.5 Implications for Future Research and Benefits

2.5.1 Potential Concept of Operations

Over the course of this effort, we have refined our vision for the OMD's concept of operations. Participating in the December 2020 SERDP Symposium was particularly valuable for drafting this vision.

In the realm of all munition-contaminated underwater sites, there is particular concern for shallow water sites where the munitions are exposed: the public is most likely to interact with these munitions. Indeed, some remediation action plans are intentionally limited to exposed munitions.

Currently, a typical remedial investigation approach involves: (1) locating targets of interest (TOI) via acoustic (or, sometimes, electromagnetic) approaches followed by (2) diver investigation of these TOI. Diver investigation is expensive, time-consuming, and risky. Furthermore, georeferencing diver-collected data is challenging.

The OMD offers a single alternative to both steps. The OMD, deployed from an unmanned surface vehicle (USV) or unmanned/autonomous underwater vehicle (UUV / AUV) can survey the site, detect TOI, and rank their probability of being munitions. OMD data (photographs and bathymetric topologies) is human readable, enabling a simple post-processing interface for quickly and intuitively verifying the TOI classification.

As the OMD is designed to operate at relatively close standoff distances (a few meters), survey speeds (area coverage rate) will likely be slower than acoustic approaches. As such, an alternative concept of operations is for the OMD to follow a traditional sonar survey. The OMD will selectively scan TOI coordinates flagged by the sonar survey. This approach has the added benefit of linking sonar and OMD data which will, over time, build a labeled sonar "image" data set – useful for training a future classifier of sonar data.

2.5.2 Recommendations for Future Research and Development

The OMD remains well suited to close operational gaps with underwater munition detection and classification. It is particularly well suited for shallow water environments (that are traditionally challenging for acoustic imaging approaches) and for exposed munitions (which are most at risk of interaction with the public). We recommend a follow-on effort to develop a fully functional prototype. This prototype should feature:

- Both structure from motion and structure light imaging. Structure from motion (and the raw images the feed the SfM algorithms) enables a broader range of automatic target recognition technologies. However, structured light imaging will work over a wider range of water conditions (turbidity), a particular concern for shallow waters.

- Improved image quality. Implementing the imaging recommendations identified on this effort should improve both raw image quality and reconstructed three-dimensional point clouds.

- Integration with an unmanned host vehicle. While the basic OMD technology is agonistic to the specific host vehicle (surface or underwater), an unmanned surface vehicle is a good candidate for the next prototype. Surface vehicles tend to be less expensive, have more flexible payload interfaces, and offer straightforward geolocation and route finding via GPS. Surface vehicles are also well suited for a number of potential test beds. ROVs, AUVs, or towed platforms would allow the OMD to be used in deeper water; no design decisions should be made that exclude eventual use on these alternate platforms.

- Automatic Target Recognition. We identified a number ATR technologies in this effort worthy of future investigation. In particular, curvature (geometry) based mapping and deep convolution neural network classification are likely the most generalizable and the most promising. A follow-on effort should continue to develop these approaches for this application.

- Simple and Intuitive User Interface. OMD data should be displayed on a user interface that features georeferencing, ranking of TOI by probability of being "munition-like", and access to the underlying images and topology for verification or rejection of the classification.

The prototype should then be validated in the open water:

- Local Testing. This testing will similar to the testing done on the initial effort. We recommend adding additional targets, conditioning the target surfaces to be more representative of years old UXO, and adding clutter targets (e.g. tires, scuba tanks, etc.).

- Dedicated UXO Test Site. SERDP is developing a number of underwater UXO test beds. In particular, the Hawaii Munitions Test Range Complex is particularly well suited for the OMD. It is a shallow water site with good water clarity, significant natural clutter, and can be readily scanned from a surface vehicle. Alternatively, sites with real and documented UXO contamination (such as Joe English Pond in New Boston, NH or the Island of Vieques, Puerto Rico).

- Other Test Sites. If necessary, additional work could include the addition of more challenging test sites (those with more turbid water or those deep enough to require integration with an undersea vehicle).

This plan should be sufficient to demonstrate the utility of the OMD as a tool for underwater munition site remediation.

3 OBJECTIVES

Many active and former United States military installations include underwater sites that are contaminated with munitions. To understand this hazard and to guide remedial action, improved survey tools are needed to assess the locations of munitions and their condition. The overall objective of this effort is to develop an Optical Munition Detector (OMD, Figure 1) that uses optical imaging techniques to detect and classify munitions underwater. The OMD is designed to close operational gaps with existing remedial investigation technologies. The overall objective of this limited scope effort was to design, fabricate, and demonstrate a proof-of-concept OMD. The specific objectives of this effort were:

- Develop a set of specifications for an optical instrument to fill SERDP's operational gaps;
- Design and fabricate a proof-of-concept demonstration system;
- Demonstrate, in representative open water conditions, the ability to scan munition-mimicking targets using data collected by the proof-of-concept system;
- Research candidate algorithms for automatic target recognition (ATR) well suited for OMD data; and
- Develop the high-level design for a fully functional prototype.

4 BACKGROUND

A major concern at munition contaminated sites is public interaction with the munitions. For underwater sites public interaction is mostly constrained to shallow water (where people are most likely to recreate) and exposed (i.e., unburied or proud) munitions (that are visible to the public).

Remediation of munition contaminated sites requires remedial investigation (RI) to determine the type, extent, and location of the contamination, and remedial action (RA) to mitigate the contamination. Effective RA is only possible with a detailed plan developed by careful RI. The vast majority of RI efforts to detect, classify, and localize underwater munitions use acoustic technologies, and not without good reason – sound propagates very well in water, can penetrate the seafloor, and there is typically reasonable acoustic contrast between munitions and the surrounding medium. Electromagnetic (e.g. using electric field distortion or magnetic anomalies) detection and classification techniques are also widely investigated.

However operational gaps remain (SERDP 2018). The precise orientation of the munition has a very large effect on the results obtained with acoustic and electromagnetic technologies. They are also challenged in the presence of clutter such as rocks and gas bubbles for acoustic approaches, and metal fragments for electromagnetic approaches. Multipath reflections (e.g. from the water column surface and bottom) can confound acoustic propagation in shallow environments.

These challenges can make detection and, in particular, classification of munitions very difficult. The outputs of these technologies, even though often called "images", are rarely intuitive for an untrained observer to interpret and are still often challenging even for trained observers. As

such, human confirmation of automatic target recognition (ATR) algorithms before RA is difficult and full of uncertainty. In many cases, particularly for exposed munitions, regions of interest flagged by acoustic and electromagnetic approaches must be visually inspected by divers. This is an expensive, risky, and time-consuming effort.

Finally, in assessing the need for RA, knowledge of the condition of the munitions is important. Are the munitions heavily corroded or otherwise structurally compromised? Are they already leaking their constituents to the environment? Or are they relatively intact and structurally sound? Existing RI approaches cannot readily answer these questions.

Our hypothesis is that optical imaging approaches will be valuable to closing these operational gaps.

5 MATERIALS AND METHODS

The Optical Munitions Detector uses optical imaging techniques to survey the bottom of the water column, extracting information (high-resolution, three-dimensional topology, optical contrast, and color) useful for munition detection and classification. The OMD has a number of advantages over traditional underwater acoustic imaging:

- It uses lower power (order tens of Watts), making it well suited for long deployments on unmanned vehicles;
- It has high resolution (sub millimeter), allowing for detailed characterization of munition condition and integrity;
- It is sensitive to optical contrast (e.g. markings and lettering) and color, features that will be very useful for ATR and human confirmation of positive ATR hits;
- It preserves the target object's geometric shape, another very useful feature for ATR and human confirmation of positive ATR hits;
- It provides immunity to multipath scattering, making it particularly well suited for shallow water missions;
- Its imaging process is less confounded by bottom clutter;
- OMD images are well suited to leverage the tremendous amount of ongoing effort of image classification via deep learning;
- Its output images are naturally and intuitively human-readable, allowing for a simple human-in-the-loop check for false positives on RI data before beginning RA;
- It is platform agnostic and can be deployed on surface vessels (manned or unmanned, autonomous or remote) or undersea vehicles.

The OMD creates three-dimensional topology via two mature and complementary optical imaging techniques: structured light and structure from motion. These techniques are introduced in more detail in the Appendix.

We divided the technical work on this project into four tasks: system specifications, proof-of-concept design and fabrication, proof-of-concept demonstration, and recommendations for a next-generation system.

6 RESULTS AND DISCUSSION

6.1 TASK 1. SYSTEM SPECIFICATIONS

The goal of this task was to develop a detailed specification table for the proof-of-concept OMD and for a future fully featured and fully functional OMD (Table 1). This specification table is a living document that will be updated as new information becomes available throughout future programs.

Table 1

Category	Characteristic	Proof-of-Concept Spec	Full System Spec
Target UXO	Type	Replica UXO, possibly including • 81mm mortar • 155mm howitzer • 25mm bushmaster • 2.75” hydra rocket	Replica UXO and/or inert rounds
	Orientation	Any	Any
	Surface Condition	Clean and known in advance	Clean, weathered, or rusted
	Extent of Burial	Largely unburied	Unburied and partially buried
	Clutter	Naturally occurring, limited	Naturally occurring and intentional mand-made clutter
Environment	Depth	<10’ (facilitates testing in lake environment)	<150 feet (No need to worry about targets that are deeper that recreational diving limits - 120 feet)
	Water Clarity	Uncontrolled	Determine standoff vs. turbidity trade
	Temperature Range	TBD	-3° to 30° C
	Ambient Light Level	Low (shadowed or night) owing to component limitations	Day or night
	Clutter	Uncontrolled; consider imaging rocks or other clutter	TBD

Host Vehicle(s)	Payload Power Requirements	Powered from laptop battery and/or deep cycle marine battery	<100-watt average - (Supports long-duration operation from battery or power supply of host vessel)
	Payload Size	Small enough to be handled by person in rowboat	TBD pending final optical design
	Payload Weight	Light enough to be handled by person in rowboat	Neutrally buoyant
Survey Parameters	Area Coverage Rates	N/A	TBD
	Probability of Detection	N/A (limited number of targets used in testing)	TBD
	False Alarm Rate	N/A (limited number of targets used in testing)	TBD
Data Processing	Processing Speed	Off-line, non-real-time	Off-line, non-real-time Possible low-overhead pre-processing
	Data Storage	Onboard storage of raw images for subsequent processing	Onboard storage of raw images for subsequent processing
	System Control	Topside control of: image type (laser scan/photo), laser on/off, and auxiliary light on/off, start/stop data collection	Remotely controlled with support for control by autonomous host vessel
	Target Classification	Detection only	Detection and classification into broad categories; ranking results by likelihood
	User Interface	Manual processing scripts	Graphical user interface displaying ranked assessment of regions of interest and sample images for target verification.

6.2 TASK 2. PROOF OF CONCEPT SYSTEM DESIGN AND FABRICATION

The goal of this task was to design and fabricate a proof-of-concept (PoC) OMD that demonstrates key system functionality.

6.2.1 System Design

We used the system specification to identify key performance parameters for the OMD optical components. In particular, resolution requirements drive camera pixel count and frame rate. Small rounds, on the order of 20mm, should be clearly resolvable in the data. Practically, we interpret this to mean their images should span at least 10-15 pixels. For similar reasons, the spatial offset between subsequent images should be on the order of 5-10mm, indicating that high speed imaging is necessary for reasonable area coverage rates.

Reasonable underwater fields of view (to keep distortion low) can go up to roughly 60°. Based on this assumption, target per-pixel resolution, and the desire to maximize scan area, we targeted a scan altitude of approximately 1.5m and a swath width on the order of 1.5-2.0m.

6.2.2 Component Selection, Acquisition, and Characterization

Cameras. For proof of concept testing we selected the FLIR GS3-U3-51S5[1] camera. This compact camera has a versatile C-mount lens attachment, high resolution (2,448×2,048) and a relatively high frame rate of 75 frames per second. We also have extensive experience with this camera family, and its USB interface allows easy and flexible integration. We have purchased the color version of this camera for the structure from motion system, and the monochrome version for the laser scanner. By pairing these cameras with a Kowa LM5JCM lens[2] (5mm focal length) we achieve a 1.7m swath width at an altitude of 1.5m. The resulting per-pixel spacing is 0.781mm, which is in line with our target resolution.

The 75Hz frame rate of this camera results in a platform speed of 0.75m/s if we assume a 10mm step between images. This rate is acceptable for Phase I testing but we will likely want to increase it in the fielded system. One factor is selecting the FLIR GS3-U3-51S5 camera is that the underlying Sony IMX250 image sensor is widely used, allowing for low-risk transition to a higher speed camera. In our initial effort the increased expense and development effort for a high-speed camera is not justified; but following proof of concept we anticipate moving to a camera such as the FLIR ORX-10G-51S5[3], which operates 162Hz.

Laser. For proof of concept testing we will use a 520nm single-line Stringray laser from Coherent. This laser is Creare-owned, available for immediate testing, and tuned for eye safety. The 520nm wavelength offers low absorption through water and a proven high-quality projected line. Again, following proof of concept, we may want to upgrade to a more powerful laser. But for initial system characterization an eye-safe laser greatly improves the ease of testing and eliminates the significant cost of safely handling a higher power laser. Further, we have previously used this laser for underwater applications and expect acceptable results based on these earlier projects.

Structure from Motion Illumination. We selected and purchased the Kraken Solar Flare Mini 12,000 (henceforth referred to simply as "The Kraken") as the structure from motion light. This light is designed for underwater use and is rated as waterproof down to 330 feet. It has a

[1] www.flir.com/products/grasshopper3-usb3/
[2] www.kowa-lenses.com/en/lm5jcm-2mp-industrial-lens-c-mount
[3] https://www.flir.com/products/oryx-10gige/?model=ORX-10G-51S5C-C

maximum output of 12,000 lumens with 80%, 60%, 40%, and 20% settings which gives us the flexibility to adjust for test conditions.

Host Vehicle. We purchased a small boat (Figure 7) to host the proof of concept OMD.

Figure 7. Photograph of the Boat Acquired to Host the Proof of Concept OMD

6.2.3 Assemble Proof of Concept System

Figure 8 shows our assembled proof-of-concept system. This system was fabricated to our design with one notable exception. As we had to rush to stay ahead of our test site freezing over, the bandpass filter for the monochrome structured light camera did not arrive in time for incorporation into our test setup.

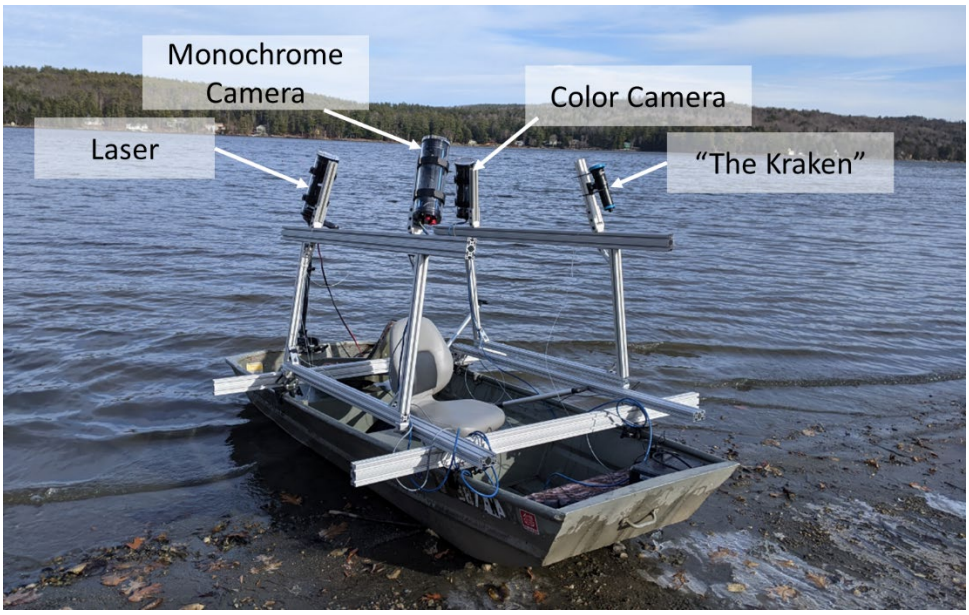


Figure 8. The Assembled Proof-of-Concept System.

6.3 TASK 3. PROOF OF CONCEPT DEMONSTRATION

The goal of this task is to evaluate the proof-of-concept instrument's performance in the open water.

6.3.1 Research and Acquire Munitions-Mimicking Targets

We spoke with underwater UXO mobility expert Dr. Joesph Calantoni of the Naval Research Laboratory and watched one of his webinars summarizing his (and others) work with underwater UXO. There appeared to a few popular choices for target UXO common to many studies. 81 mm mortars and 155 mm Howitzer rounds are particularly common. 25 mm chain gun rounds, 105 mm artillery rounds, and Hydra rockets are also used. For consistency with the existing literature, we selected our targets from these. As our system is sensitive only to shape, size, color, and texture, we purchased inert replicas from an online source (Inert Products LLC <inertproducts.com>). **Figure 9** shows the inert munitions purchased in support of this effort: a 105mm Artillery Projectile, two 25mm Bushmaster Rounds, a 81mm Mortar, a 155mm Artillery Projectile, a 81mm Mortar (with UXO texture and coloring), and a 105mm Artillery Projectile (with UXO texture and coloring).

Figure 9. Inert Replica Munitions, from Left to Right: 105mm Artillery Projectile, 2x 25mm Bushmaster Rounds, 81mm Mortar, 155mm Artillery Projectile, 81mm Mortar (UXO), 105mm Artillery Projectile (UXO).

6.3.2 Develop Test Plan

We developed a high-level draft test plan for the proof-of-concept OMD. The test area will be (roughly) 10 feet wide and 20 feet long at an approximate depth of 6 feet. The targets from task 3.3.1 will be lowered into the test area using attached retrieval ropes. We will then make multiple passes over the test area with the proof-of-concept OMD from different directions varying the light

brightness and vessel speed to gather multiple data sets for comparison. This plan was adapted on the fly during our first test event and then again between our first and second events as described below.

6.3.3 Execute Proof of Concept Tests

First Test Event. Our first test event took place at Mascoma Lake (Enfield, NH) on 19 November 2020. The targets were secured via 300 lb. test fishing line to anchors which were in turn secured to buoys with rope (Figure 10). The targets were deployed in an approximately 10-foot by 20-foot area at a 6-foot depth and all the anchors were pulled to the shallow side of the test area when the targets were deployed. During testing we were only able to travel parallel to shore due in order to avoid running into the buoys. The bottom at this location was full of vegetation (Figure 11).

We started testing in late morning. Images from this first run were somewhat blurry and overexposed (Figure 11, Left) so for the second run we reduced the camera exposure. The second batch of images were clearer and not over saturated (Figure 11, Right). However, there was still too much ambient lighting to resolve the laser line in the structured light images. This was due mostly to our very low-power (but eye safe) laser projector and the lack of the bandpass optical filter.

Through the first day of testing we identified multiple areas with room for improvement for future testing. Some of the targets sank slowly so additional holes were drilled to facilitate faster sinking. The buoy-based deployment system was the weakest link: it limited our approach paths to the target area and made judging the location of the target area difficult as the buoys shifted in the wind. The buoy lines were also easily entangled with our test platform. Lastly, there was too much ambient light during the daytime for our (under-powered proof-of-concept) structured light imaging system.

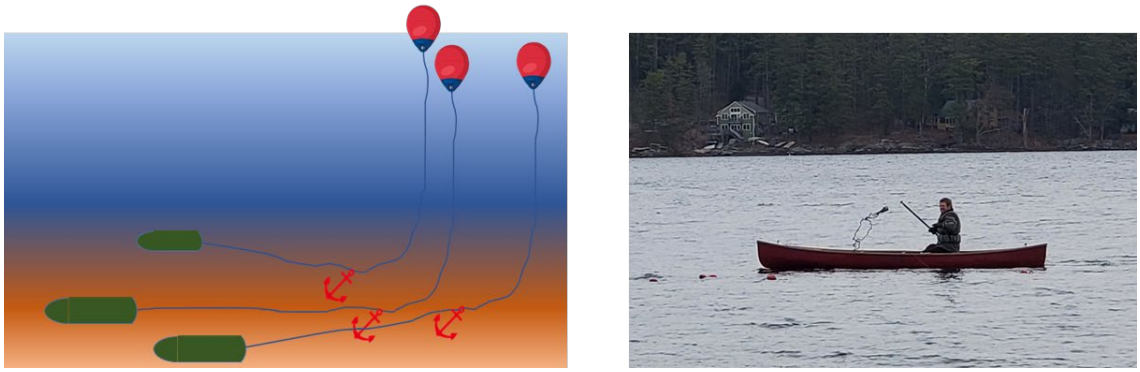


Figure 10. Test Target Setup (Left) and Target Deployment (Right)

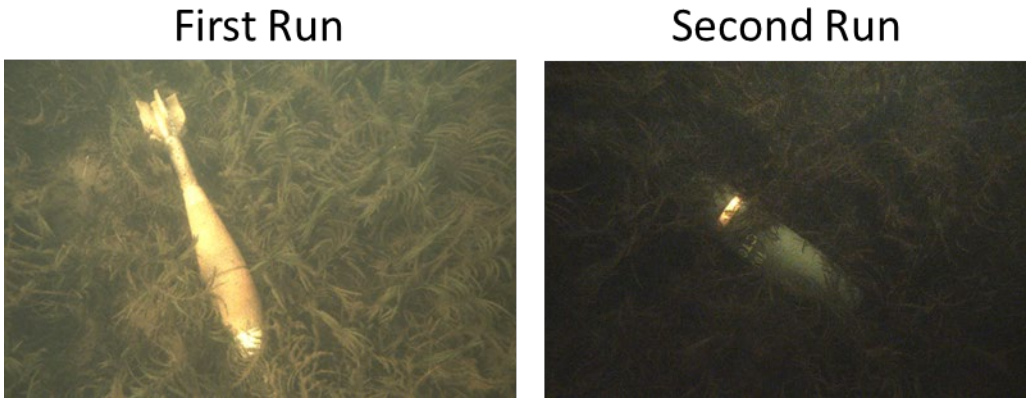


Figure 11. Selected Color Images from the First Test Event.

Second Test Event. Correcting our test setup and plans for the lessons learned during the first test event, we performed a second test event on 3 December 2020 at the same location. (We were fortunate that the lake froze over much later than usual this year.) Overall testing went much smoother and resulted in much higher quality data.

We ran 200 lb. test fishing line from the targets to the shore (Figure 12, left). There were no buoys or buoy lines to entangle with the test craft. This greatly simplified the testing. The previously troublesome targets sank quickly with the added vent holes. Prior to deploying the targets, we marked the four corners of the test range (Figure 12, right) with flexible poles that would not shift position with the wind. We secured the poles to their bases with springs so they would not be tipped over when struck. The poles had reflective tape at 6-inch intervals and battery powered LED glowsticks mounted to the tops to help with visibility at night. The regular spacing of the reflective tape also allowed us to judge the depth of the test site. We were able to easily layout and adjust the target area as well as approach from multiple angles.

Our test setup went so smoothly that we were able to begin testing during daylight. Images were of good quality (Figure 13, left) but there were some optical caustics caused by the low angle of the late autumn sun that would interfere with the image processing. Thus, we waited until the sun had set behind the hills for the second test run. With little to no ambient lighting, the “Kraken” was the main light source for the SfM images (Figure 13, right), allowing us to evaluate its optical performance and our geometric setup (camera to light source spacing and angle). We were also able to resolve the laser line in the structured light images.

After testing at the heavily vegetated six-foot-deep site, we reset the targets at a shallower site (4 feet) with less vegetation. This site had a couple cm of mud above sand.

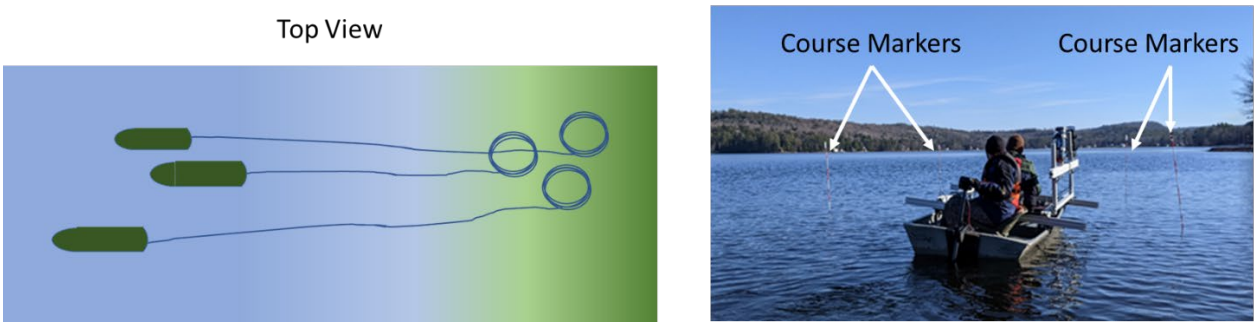


Figure 12. Test Layout During the Second Test Event.

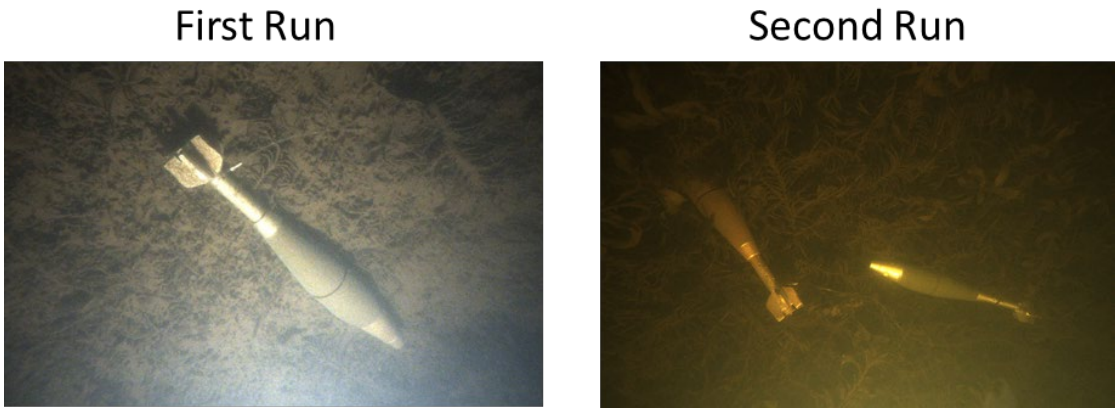


Figure 13. Selected Color Images from the Second Test Run.

6.3.4 Process Test Data

We limited our data processing to data collected during the second test event (the first test event being written off a learning exercise).

System Calibration. We calibrated both the structured light and structure from motion systems. The cameras were calibrated in their housings, but in air, using standard techniques. Specifically, a calibration target was imaged in a variety of poses and an optimization used to estimate both the parameters of a pinhole camera model and the poses of the target. We used custom code that is based on tools available in the OpenCV (opencv.org) software package. Following this in-air calibration, a "pinax" camera model[4] was constructed to account for refraction through the housing window and into the surrounding water. We have previously validated this approach for the calibration of underwater cameras using more convenient in-air laboratory data and were able to also apply the approach successfully to the OMD imagers. Calibration of the laser line projector in the structured-light sub-system consists of defining the laser pose within the camera coordinate frame. This was done using as-designed mechanical mounting parameters.

[4] Łuczyński T., Pfingsthorn, M, Birk, A., "The Pinax-model for accurate and efficient refraction correction of underwater cameras in flat-plane housings," Ocean Engineering, Vol. 133, 2017, pp 9-22.

Structure from Motion Subsystem. Figure 14 shows an example structure from motion (SfM) scene from one data set. This data set applied the camera calibration to produce results in engineering units (e.g. meters). The four munitions in the scene are clearly visible and properly shaped. Vegetation is also clearly visible in the meshed (bottom) scene.

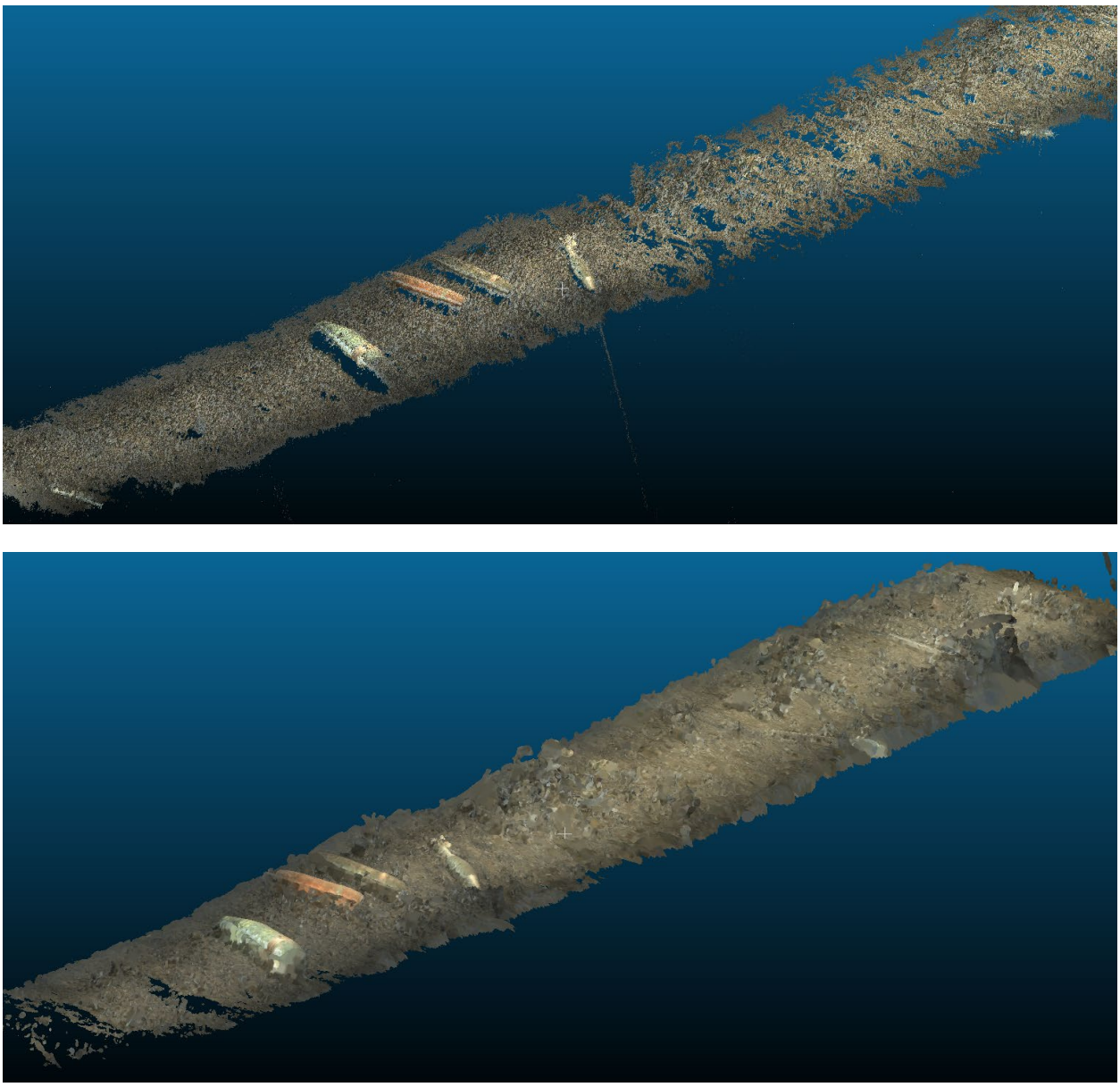


Figure 14. Example Reconstructed Structure from Motion Scenes from an Entire Imaging Run. Top: point cloud; Bottom: meshed point cloud.

Structured Light Processing. To illustrate the structured light data processing process, we performed some initial data exploration and processing for the structured light imagery. Specifically, we confirmed that useable laser-line data was collected and that our algorithms could automatically extract an accurate line profile from the images. While the low power of our

proof-of-concept laser precluded daytime data collection, images collected at night exhibited good signal-to-background contrast and low noise. We expect to achieve similar images during daylight in a next-generation system, by using narrowband optical filter (excluding wavelengths away from the laser illumination) and a significantly brighter laser.

Routine parameter tuning applied to existing custom algorithms resulted in clear laser-line segmentation, as shown in Figure 15. Laser line extraction is the first algorithmic step in the structured-light data processing.

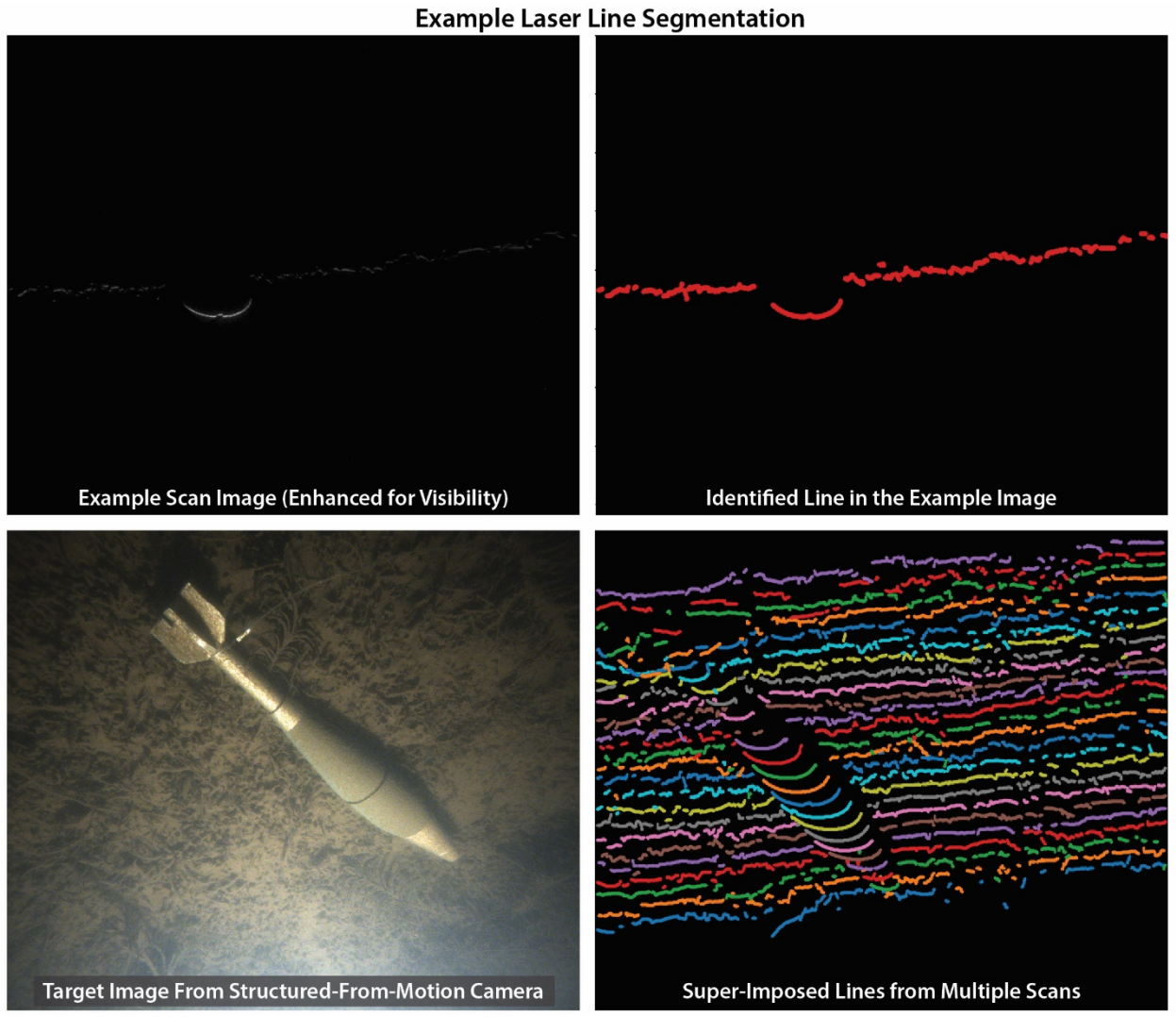


Figure 15. An Illustration of Successful Laser-Line Segmentation in Structured Light Test Data. The upper left panel shows an example laser scan image; filtering and intensity-level adjustments have been applied to aid in visibility. The upper right panel shows the pixels identified as lying on the laser line. The lower left panel shows an image, from the structure-from-motion system, that captures the model munition being imaged. And multiple identified laser lines (each from a separate image) are offset and overlaid in the lower right panel. This combined image illustrates that the munition is sampled by the structured-light imager, and that the height-dependence of the laser line position is clearly captured.

Point Cloud Construction. We synthesized 3D measurements of our lake test scene from the structured light data. This effort builds on the laser-line segmentation described above by using the identified pixels in conjunction with the calibrated system model to construct a measurement through triangulation. Each image in the dataset gives a line of three-dimensional points, which are then concatenated in a manner consistent with the estimated speed of our test vessel. The result from one pass over our test field is shown in Figure 15. The UXO targets present in the field of view are clearly visible in the data, and the structured light imaging approach results in a quantitative characterization of their size, shape, and orientation.

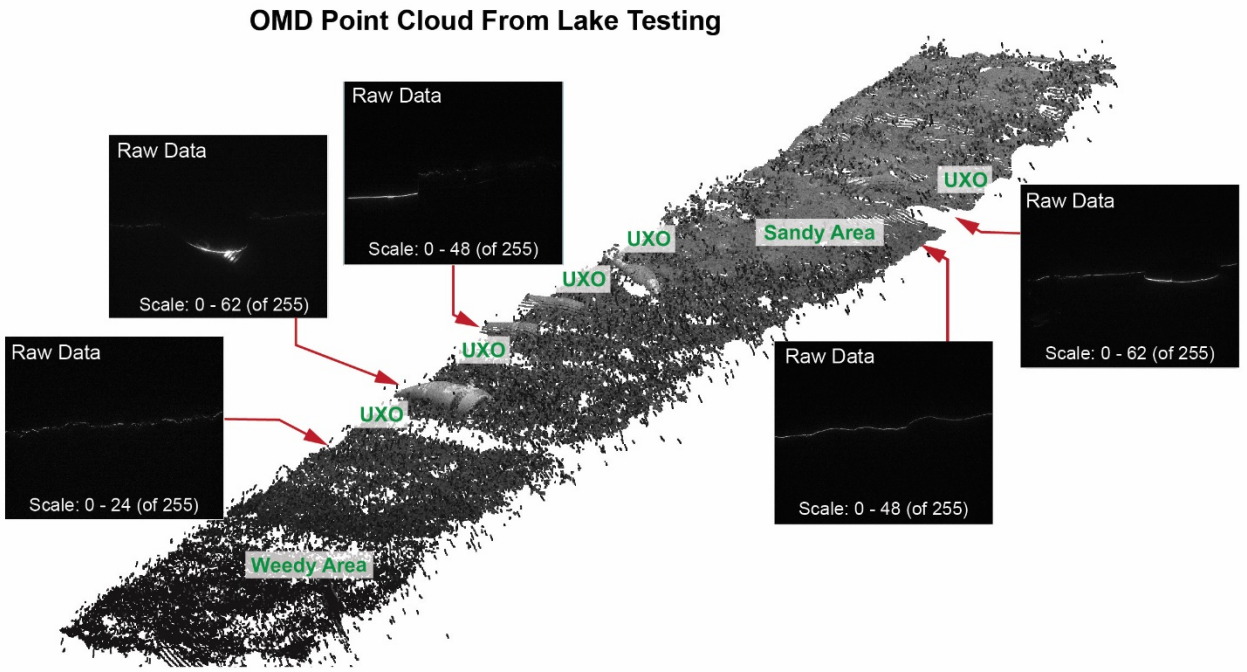


Figure 16. A Three-Dimensional Point Cloud Measurement of Our Test Scene. The point cloud is rendered in the center of the plot and example raw OMD images are shown at select locations. Shading of data points within the point cloud are based on the observed brightness of the laser line in the imagery, and thus have a correspondence with the reflectivity of the target. Points of interest within the scan, including 5 observed UXO targets, are labelled. Note that the raw images have been scaled to display over a narrower range than the full 8-bit acquisition to enhance visibility.

6.4 Design of Next-Generation System

The goal of this task is to develop a high-level design of a more mission representative next-generation system. Our work focused on two main areas: improving image quality and identifying candidate techniques for automatic target detection and recognition.

6.4.1 Optical Imaging Improvements

We asked our expert underwater imaging consultant, Dr. Jules Jaffe of the Scripps Institute of Oceanography, to review our imaging design and sample data and to provide us with both general and OMD-specific guidance for improving our image quality. As part of this work, Dr. Jaffe also prepared a concise underwater imaging primer, included as an Appendix (Section 9.1) to this report. Below we include Dr. Jaffe's recommendations for the next-generation OMD.

Overall Considerations for the OMD System. For the most part, the optimal design of underwater imaging systems entails the choice of both illumination lighting and also the geometry of the placement of components, here, being the camera and the lights. In this section we consider the criteria for optimal design in the context of the experiments that were performed by Creare and recommendations for future configurations. Summarily, it is noted that the laser sheet system will typically be power limited, not backscatter limited and, will therefore obtain images in situations that would not allow the SfM system to work. This may be the "go to" option in very turbid water but, clearly, has limitations in very turbid situations.

As discussed above, the SfM system that relies on broadly illuminated areas will potentially suffer from excessive backscatter. Ways to eliminate this consist of either using underwater lights that have narrower beam patterns (mostly desired) or using physical blocks of the beam to narrow down the illumination. The latter is more versatile but risks the presence of internal reflections while the former conserves energy. Simply put, why illuminate an area that is wider than the field of view of the camera?

Specific Considerations for Laser Light Stripe Imaging. As only one wavelength of laser light is being used, an optimal wavelength of 532 nm would be a good choice. In addition, images that include both backscatter as well as altitude would be a full use of the data. Creare has extensive experience in using laser sheet imaging from our previous work that resulted in excellent images. Ideally, the beam should illuminate a width close to the single pixels of the camera, thus maximizing resolution. In addition, the leading edge of the beam can be used instead of the maximal value. I also propose that Creare consider the dual use of combining reflectivity images with the laser sheet. The idea is to propagate a laser sheet and also use the broader field illumination of their Kraken system. That can be superimposed together. In this case, in case the SfM is not working, it might be possible to obtain elevation information and surface reflectance w/o the use of the SfM.

Specific Considerations for the Broader Illumination System for Structure from Motion. Consideration in the use of SfM systems for underwater imaging raises a number of interesting issues that have, perhaps, not been fully considered in the journal literature. Considering the literature from 2020, van Scheltinga (2020) considered a lab system to reconstruct underwater dunes with good results. The unique advantages were that they could use the image "texture", that is the color and intensity changes in the images, to perform the necessary pose and range estimation for SfM to work. In another recent article, Hatcher et al. (2020) employed the use of the SQUID-5 towed surface vehicle to perform multi-view imaging in the Florida Keys at depths of 3 m to 9 m with mm to cm resolution over several sets of surveys with excellent repeatability of 3 cm RSME. The system employed the use of 5 MP FLIR cameras (similar to the cameras used in the PoC OMD) with a fixed length 6 m focal length lens. As documented, the article used only passive solar illumination and the results were constrained by both sea state and the turbidity of the water. Presumably, only relatively clear water conditions facilitated the system's success. In addition, the use of solar illumination would allow the general reflectivity maps of the bottom terrain to be somewhat uniformly lit with almost no shadows. This is one of the main advantages of diffuse solar illumination as the presence of strong shadows due to surface roughness might confuse the SfM algorithm. In addition, the article contains an authoritative bibliography of the use of SfM in marine habitats. We also note that this area has been growing rapidly, due to the ease of software

implementation and the success of the method that requires little navigation and no inherent information about camera pose.

General Advice for the Next-Generation System. As is clear from the above discussion, optimizing performance strongly depends on the environmental parameters related to absorption and scattering. We also note that these parameters may be very local and that complicates matters greatly. As one example, the "nephloid layer" that many bottom environments are subject can be a relatively thin or thick, representing an extreme vertical dependence on absorption and scatter. Optimizing for environments that are spatially varying presents interesting challenges for underwater imaging that have only been minimally addressed. As almost every underwater imaging professional will validate, the use of in situ data that addresses the scattering and absorption properties of the water are a prerequisite to system design and deployment. Perhaps a simple Secchi disk that is color dependent could be used (and invented!!) to allow simple water properties to be measured. A more sophisticated approach would employ a multi-wavelength transmissometer. The Jaffe lab has used a 532 transmissometer in many applications. We repeat that optical image system configuration can only be performed with knowledge of environmental parameters.

Advice for the Next-Generation System: Laser Light Sheet. As discussed above, the laser light sheet will likely yield the best images in the most turbid situations. In this case it makes sense to use a powerful laser that has a narrow divergence angle. In our own laser lidar imaging system we used a pulsed 3W laser at 532 nm. Lower power lasers will likely work in environments where the water has less absorption. In addition, we note that for lake imaging, a wavelength in the orange to red would likely have better penetration capability as long as the cameras are sensitive at these wavelengths. The other concern in structured lighting is the angle between the source and receiver. As Figure 35 (Appendix) clearly shows, the height is calculated via the displacement of the reflected light beam. As such, maximizing the angle between source and receiver will yield the best range estimation. However, this also yields more attenuation as the light needs to travel a longer distance. Therefore, given anticipated resolution needed for the task, an angle should be determined that is adequate. If this requires a source-receiver separation that is beyond the pragmatic use of the platform, the maximal distance is a desirable option.

One idea that merits consideration is the dual use of a light sheet in concert with a traditional imaging camera, such as that being used in the SfM configuration. Only here, a 3D point cloud is configured that uses the reflectance map of the color image due to a broad spectral illumination that is then combined with the light sheet depth estimate. Figure 17 illustrates the concept. The advantage being that the image can be indexed in 3D without SfM concerns.

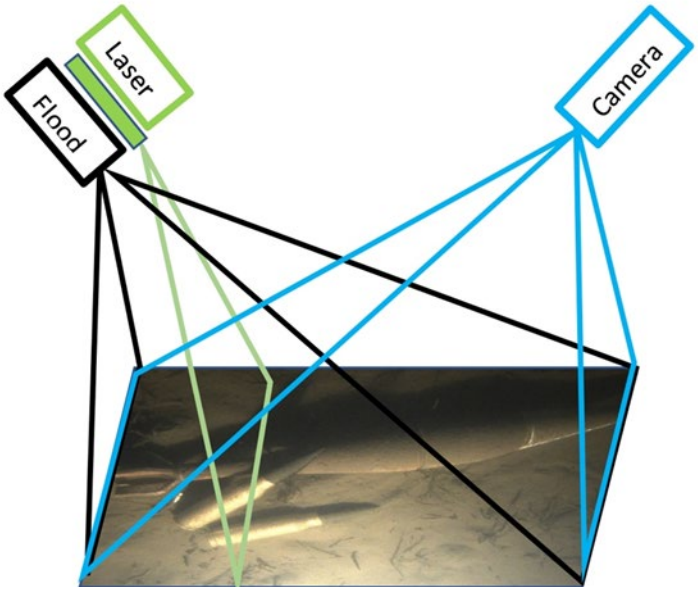


Figure 17. A Possible Configuration that Combines Light Sheet and Reflectance Imaging.

Advice for the Next-Generation System: Floodlit Scene for SfM. In considering the use of SfM for the camera images that have been flood lit, the major concerns are whether: (1) There is enough contrast to actually see the images and (2) That there are enough features in the image reflectance pattern (often called by this community "texture") to be able to decide both pose (camera subject orientation) and location. Together, these two features allow the SfM algorithms to compute the 3D distances of objects that then leads to the synthesis of a 3D point cloud. In consideration of (1), the necessity of minimizing backscatter, especially in turbid environments is the usual mandate for acquiring reasonable images. Here, we note that given a "standoff" distance of 2 meters and a field of view of 1.2 meters, the angle of view would be 30 degrees. Given that the strobe has a half angle of 60 degrees, the illumination is much too wide for this field of view at this given range. This likely necessitated turning the floodlight to a more downward position which wasted light. A natural solution would be to use strobes that have a narrower beam pattern.

An additional important feature of the SfM algorithms is that they are susceptible to excessive shadowing. Work in the Jaffe lab in past years highlighted the problems with SfM in that subjects that had lots of 3d surface roughness more often, than not, failed to allow the SfM system to succeed. Looking into the literature on marine SfM or SfM in turbid underwater environments, this author could not find evidence of published work in this area. Noting the article by Hatcher et al. (2020) was successful in clear water, it is likely that their illumination was diffuse, thereby facilitating the use of SfM in these environments. In consideration of this, my recommendation is that the authors employ a dual illumination system that will reduce shadows. We note that the same geometry can also be mimicked with a wide beam and several cameras that record images from a variety of camera angle beam patterns. The idea is encapsulated in Figure 18. Importantly, the geometry will need to be adjusted to minimize backscatter. So, this should be considered as a preliminary idea.

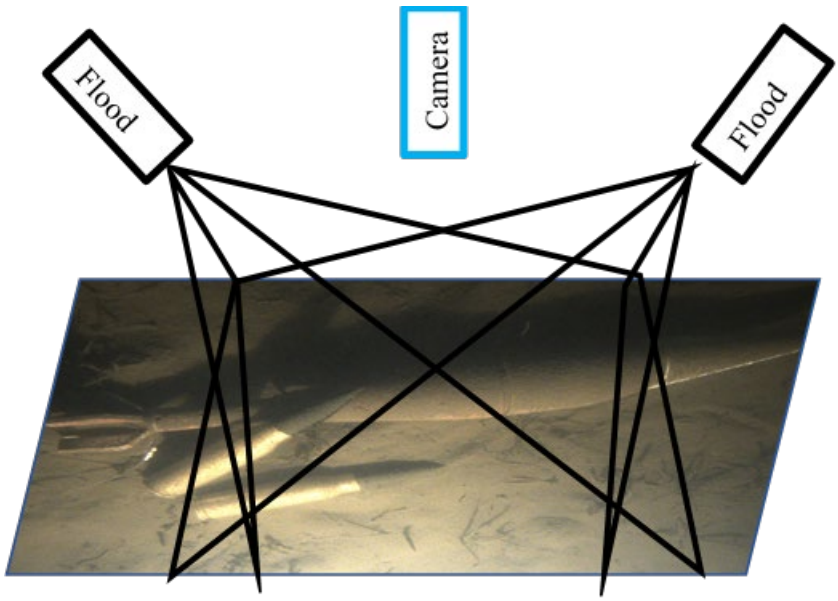


Figure 18. A Concept for a "Dual Illumination" System to Reduce Shadowing that Will Confuse the SfM Algorithm.

Dr. Jaffe's Summary. Although the preliminary data from the prototype OMD system looks good, it is likely that future work will encounter water that has higher absorption and scatter than those present in the CREARE study. In this document, a variety of issues, some novel solutions, and considerations for future work in this area are considered.

6.4.2 Automatic Target Detection: Structure from Motion Data

Geometry-Based Target Detection in SfM Data. We developed a preliminary target-detection algorithm based completely on the geometry from the SfM point clouds. By relying only on geometry, this algorithm is robust to changing light conditions, vegetation, and other underwater clutters. It takes advantage of the near constant curvature of targets over their extents. As shown in Figure 19, our algorithm has four steps: (1) compute the curvature of each points; (2) threshold the curvature; (3) connect regions with curvature that satisfies the thresholds; (4) reject regions with the wrong shape and size to produce labels of potential targets.

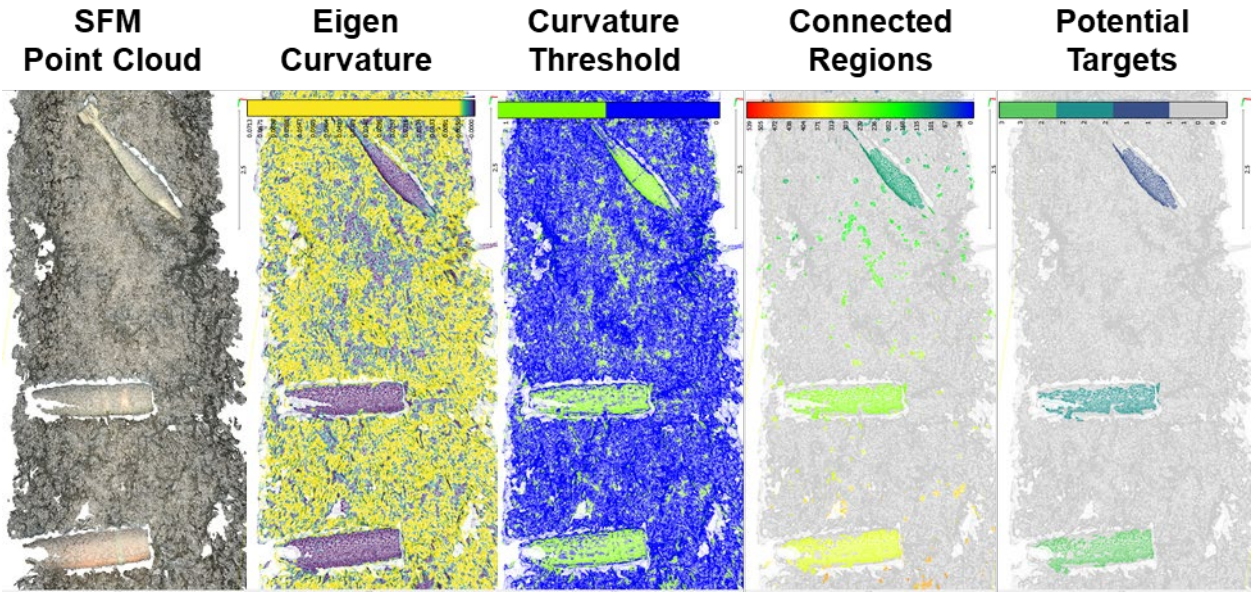


Figure 19. Detection Algorithm Based on SfM Point Cloud Geometry.

We tested this algorithm on three different cases, as shown in Figure 20. In all cases, our algorithm correctly identified all the targets (no false negatives). However, a few regions near the boundary was identified as target (false positives). Each case used the same algorithm parameters, which were tuned on the higher density point cloud. Since this particular, early SfM solution used an uncalibrated camera for these tests, we could not use physical length scales to help with the filtering. Instead, our algorithm parameters assumed on a similar density of points. These parameters were tuned using the highest density point cloud (case 2). As such, it is surprising and encouraging to the algorithm performing well on all three cases, even though cases 1 and 3 have much lower density of points compared to case 2. With a calibrated camera and additional filtering steps, it is feasible to develop a target detection algorithm using geometry data from the point cloud data alone.

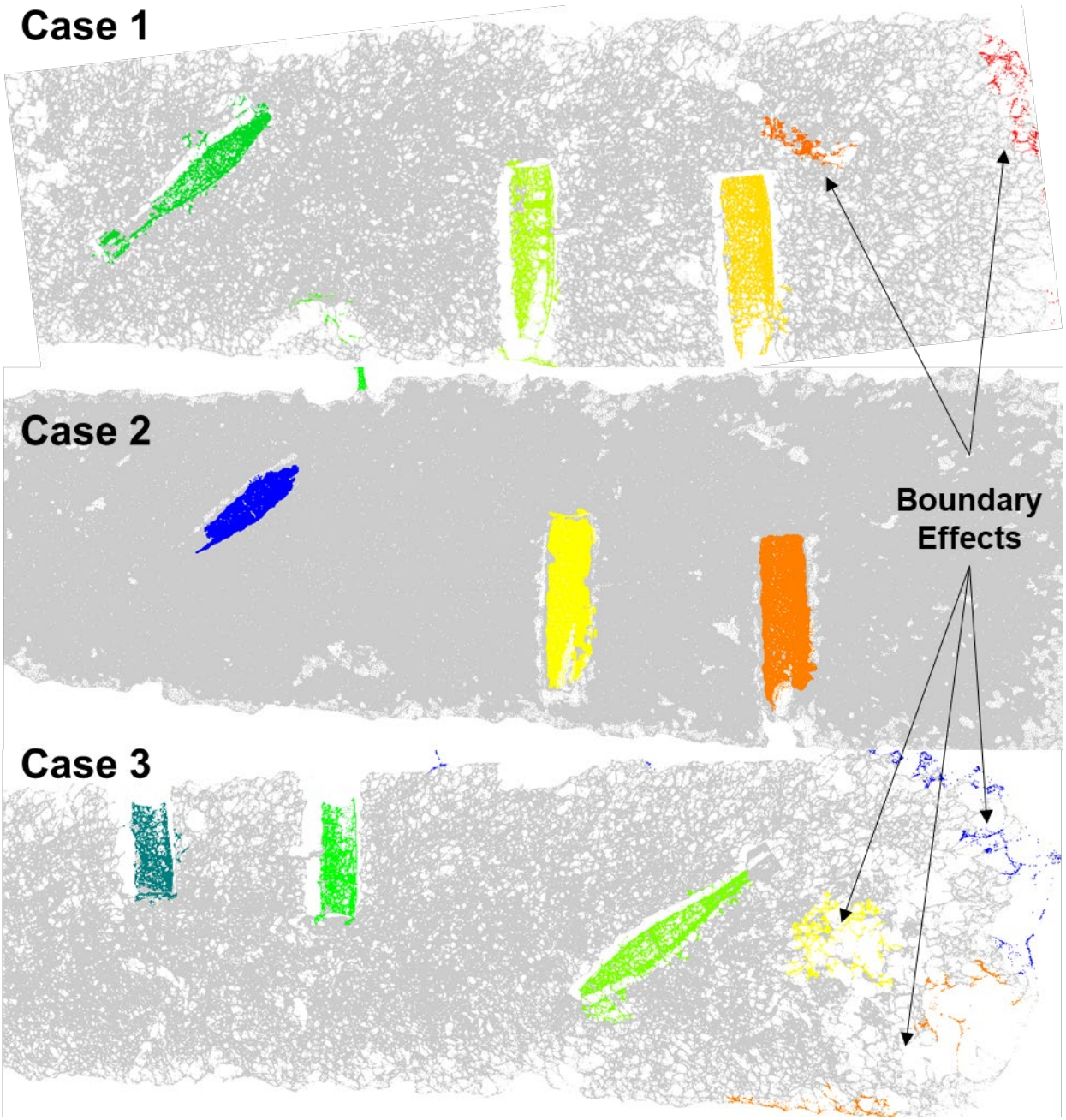


Figure 20. Detection Algorithm for Three Different Cases. In all cases, the targets are correctly identified. False positives are also identified as a result of a boundary effect, or where the point cloud is sparse. All cases use the same algorithm parameters, which suggests a carefully designed algorithm using this approach would be generally applicable.

Detailed Algorithm Description. Our algorithm starts by computing the curvature. For this step, we first smooth the local data by iteratively replacing each point by the mean of its 32 nearest neighbors. To efficiently find nearest neighbors, we organize the data into a KD-Tree data structures. Then, for each point we compute the 3x3 covariance matrix of the neighboring points as $\boldsymbol{C}_{jk} = \sum_{i=0}^{N}(\boldsymbol{X_{ij}} - \bar{\boldsymbol{x}}_j)(\boldsymbol{X_{ik}} - \bar{\boldsymbol{x}}_k)$, where $\boldsymbol{C}_{jk}$ is the covariance matrix, $\boldsymbol{X_{ij}}$ is the i^{th} neighbor's

j^{th} dimension, and N is the number of neighbors. We find the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ and compute the curvature as $c = \lambda_2 / \sum_i \lambda_i$.

Next our algorithm thresholds the computed curvature data to drop regions with high curvature. In practice, these upper and lower thresholds would correspond to the actual curvature of the targets. After this step, most of the points are eliminated, but small pockets of points remain that do not represent targets. The third step filters unconnected regions. This is done by starting with an active region, looking at its neighbors, and marking all of these with the same label until none of the neighbors are active. In regions with sparse point cloud density, the neighbor connection can span over a large distance. This can be corrected by only looking within a certain radius of each point instead of using N neighbors. The algorithm continues by looking at the next unlabeled but active region and repeating the process until all active points have a unique label.

Finally, each regions of points are filtered based on their total extents. For this part of the algorithm, we use a heuristic length scale tuned to fit our data. When the real scale of the data is known, this length scale would correspond to actual target sizes. Additionally, for further improvement, actual models of the targets could be fit to each region's data, further improving the detection accuracy.

Summary. Overall, this approach seems to work well with our Phase I data and shows feasibility of developing such a detection algorithm during a follow-on effort. It also demonstrates that the SfM geometry data truly contains information that can be used to automatically detect targets.

6.4.3 Automatic Target Detection – "Texture" of Raw Camera Images

Working with the raw camera images used to create the structure from motion data inspired us to investigate techniques to detect munitions from the raw images used to create the structure from motion point clouds.

One straightforward approach involves 'texture analysis'. Texture analysis relies on variation in image intensity computed over small regions of the image. This method was selected to assist in identification of the edges of munitions lying on the bottom as these edges appear to be regions of locally high contrast in the image. Subsequent processing could be applied to look for straight or nearly straight lines that are typical of the sides of shell casings.

Texture was computed by first converting the color image to a grayscale intensity image. Next, a texture metric was computed in a 50-pixel circular neighborhood throughout the image. The texture metric does a good job finding the regions of largest variation in image intensity (Figure 21). The bottom image showing the texture results was false colored so that the regions of highest texture/intensity variation are shown in red.

Summary. While these preliminary results are encouraging, we do not have enough data to know if this approach is generalizable to "real-world" conditions with increased clutter or weathered munition surfaces. That said, the technique is simple and easy enough to execute that it warrants investigation on future data sets.

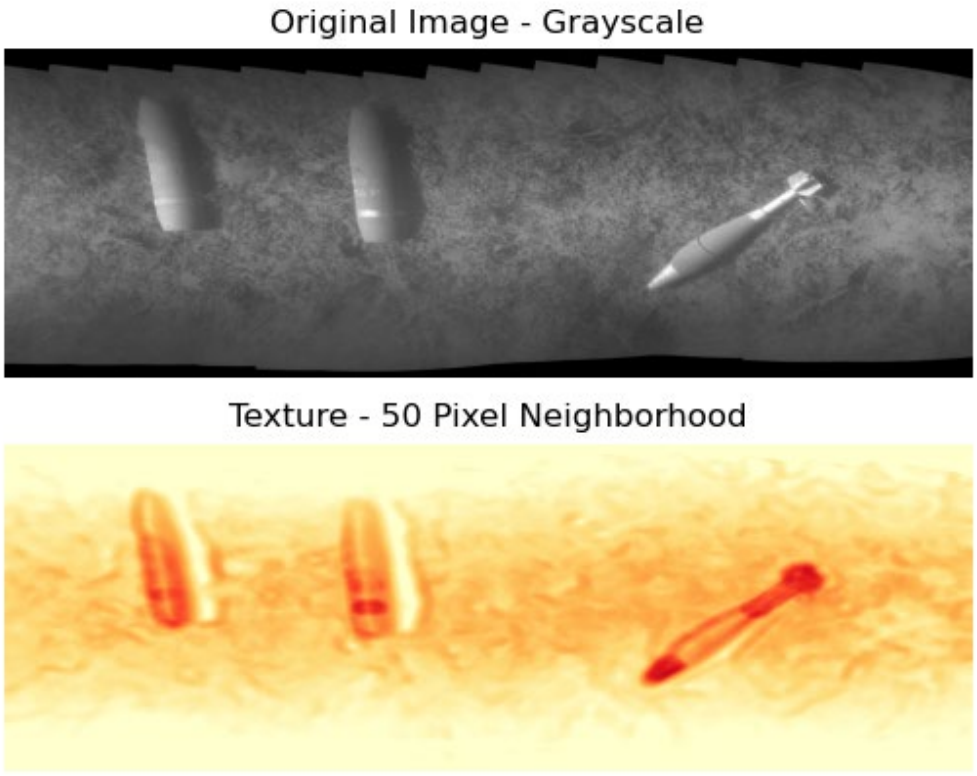


Figure 21. Texture Analysis. Image texture is computed from the grayscale intensity image shown in the top panel. The texture metric is designed to highlight regions of highest intensity variation in the underlying image (shown in red). These regions coincide with the edges of the munitions

6.4.4 Automatic Target Detection – Machine Learning on Raw Camera Images

We applied an image-based Machine Learning algorithm to detect the inert munition targets in data collected during our system tests. Specifically, we used a pretrained Deep Convolution Neural Network (DCNN) image classifier developed for a project supported by U.S. Army Corps of Engineers. Importantly, this classifier was NOT trained on underwater images or images with munitions.

For the Army effort we trained a DCNN capable of filtering images (taken with a smart phone) relevant to crowd sourced weather data collection. Essentially, this DCNN was trained to distinguish between images taken outdoors of natural scenes where weather related phenomena might be present, and images taken indoors of mostly artificial/man-made objects in climate-controlled environments. As such we theorized that the DCNN could be used to identify munitions (artificial manufactured objects that would more closely resemble objects found indoors) amongst natural surroundings (e.g. vegetation, sand, dirt – objects that the classifier would be more likely to label as “weather-related”). The training process is discussed in more detail below.

Artificial Object DCNN Classifier. The *Weather/Not-Weather* classifier was developed using supervised machine learning techniques and the *Openimages*[5] dataset was used as a source of labeled images for model training. We identified 23 weather specific categories of labeled images in the *Openimages* dataset, and randomly selected 400 images for each label. In addition, we identified 63 *Openimages* labels that could indicate being outdoors. We included 200 randomly selected images for each of these labels in the dataset representing weather-related images to be more inclusive of images that did not necessarily fit our specific weather labels but were still potentially weather related. To represent the non-weather class, we selected 200 randomly chosen images from each of 27 labels indicating an image taken indoors. The specific image categories used can be found in Figure 22.

To train the DCNN, we started with an Inception V3 model that had been previously trained on ImageNet to classify images. We trained a new two class softmax classification layer with the *Weather/Not-Weather* data, leaving the weights in the convolution layers as is. This reduces the computational cost of training and can help with overfitting in some cases. To further reduce training costs, the output from the convolution layers was precomputed for each image in the dataset so that each image could be used for multiple optimization iterations of the weights without the need to recalculate the outputs from the convolution layers.

[5] <https://storage.googleapis.com/openimages/web/index.html>

WEATHER	OUTDOORS			INDOORS	
1. cloud	1. forest	24. dirt road	47. pedestrian	1. ceiling	24. interior design
2. cumulus	2. ocean	25. trail	48. pedestrian crossing	2. bedroom	25. studio
3. rain	3. desert	26. park	49. stop sign	3. bathroom	26. recording studio
4. blizzard	4. river	27. sunrise	50. fence	4. kitchen	27. mattress
5. freezing	5. lake	28. sunset	51. deck	5. dining room	
6. drizzle	6. pond	29. beach	52. patio	6. living room	
7. frost	7. stream	30. sunlight	53. yard	7. basement	
8. sea ice	8. mountain	31. moon	54. courtyard	8. attic	
9. fog	9. prairie	32. moonlight	55. backyard	9. classroom	
10. mist	10. steppe	33. wildlife	56. playground	10. hallway	
11. thunderstorm	11. arctic	34. autumn	57. outdoor bench	11. bookcase	
12. winter storm	12. jungle	35. summer	58. outdoor furniture	12. dinner	
13. lightning	13. hill	36. winter	59. outdoor play equipment	13. bed frame	
14. dust	14. field	37. spring	60. outdoor power equipment	14. television	
15. flood	15. garden	38. roof	61. outdoor recreation	15. desk	
16. wildfire	16. farm	39. boat	62. outdoor structure	16. office chair	
17. ice	17. tree	40. skiing	63. outdoor table	17. lampshade	
18. tornado	18. city	41. hiking		18. bathtub	
19. landslide	19. skyscraper	42. camping		19. ceiling fan	
20. sun	20. cityscape	43. picnic		20. curtain	
21. sunlight	21. street	44. wildflower		21. closet	
22. sky	22. road	45. sidewalk		22. dresser	
23. precipitation	23. highway	46. boardwalk		23. wardrobe	

Figure 22. Openimages Labels Used to Train the Weather/Not-Weather DCNN. To represent the Weather class, we used 23 labels that were specifically weather-related concepts and 63 labels representing outdoor concepts. The latter allow for inclusion of images that do not necessarily fall under the specific weather labels, but could be relevant, nonetheless. To represent the Not-Weather class, we selected 27 labels likely to be associated with images taken indoors. Openimages allows for distinct images to be associated with multiple labels. For the Weather class, we allowed images that might also be associated with indoor labels. For the Not-Weather class we only selected images that were not associated with the weather labels or the outdoor labels.

Application to Munition Target Detection. We tested the DCNN on images collected for SfM from three runs ("5-11", "5-12", and "5-13") with 246, 119, and 100 images respectively. The 5-11 run included images of three distinct munitions of different shapes and colors with and without visible text markings, lightly to moderately covered in sand with some reflective glare from the illumination source. The background was a mix of dirt/sand and vegetation. 5-12 showed images of a single munition with text markings, also lightly covered in dirt/sand. Background was the same mix of vegetation and dirt/sand. More glare and shadows were present from the illumination source in this run. 5-13 included a mortar (with simulated UXO conditioning) and the two 25 mm rounds with a background more dominated by sand/dirt then vegetation.

The FLIR GS3-U3-51S5 camera captures high resolution (2,448×2,048) color images. As it is built on a pretrained Inception V3 architecture, the existing DCNN requires color (RGB) input images that are a resolution of 299×299 pixels. There are two options for processing the SfM images in order to address the resolution difference: (1) resize the images, (2) select crops from the images. We found better performance by selecting crops from the images rather than resizing. One possible explanation for this is that the image scale was such that in most, if not all, of the SfM images a munition was always surrounded by background vegetation/soil/sand. Selecting small patches tended to result in crops that were dominated more by either a munition or by

background and not as a heterogeneous as the native resolution images. Figure 23 shows an example of this – the crop with clear munitions is has a much higher probability of being "man-made" than the crop with just the tip of a round present.

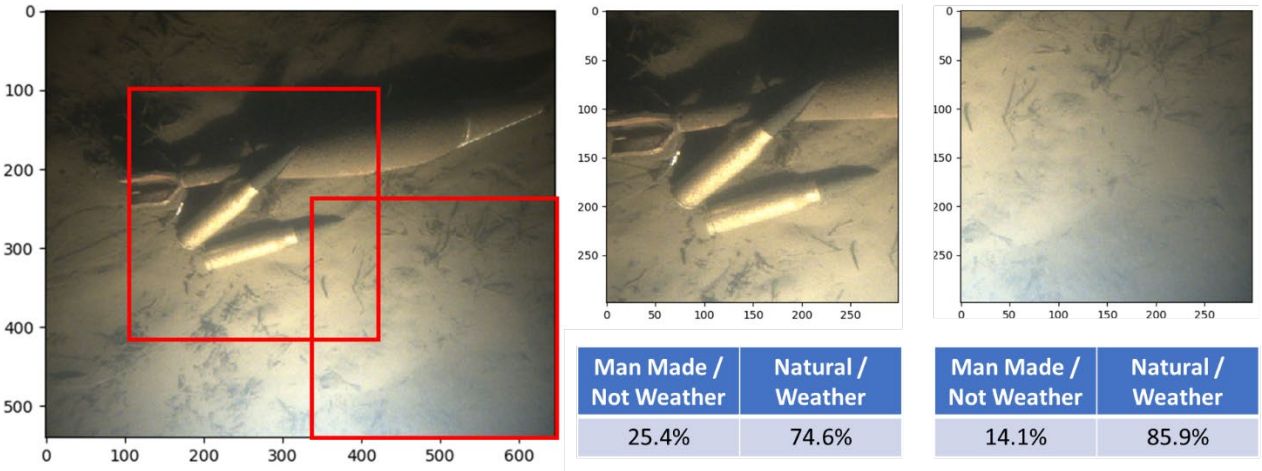


Figure 23. Two Example Crops from a Single Image and Corresponding DCNN Probabilistic Output. The image with clear munitions (Middle) has a much higher probability of being "man-made" than the crop with only a small portion of a munition present (Right).

To illustrate the effectiveness of image crops we "slid" a crop pixelwise across an image panorama, calculating the probability of being "man-made" at the center pixel. Figure 24 overlays the classifier probability on the original image. The munitions in the scene have a much higher likelihood of being "man-made" than the surrounding vegetation.

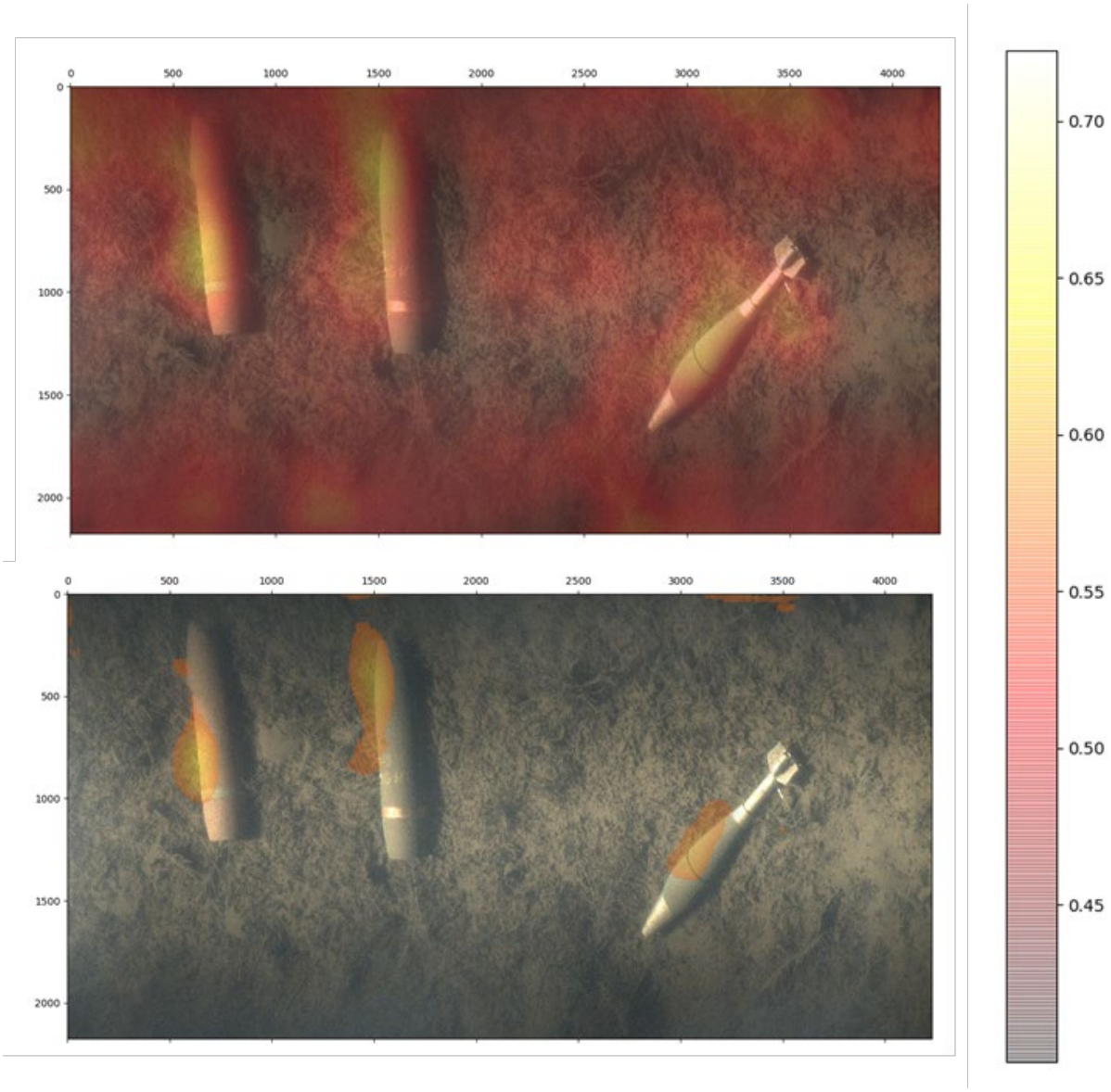


Figure 24. DCNN Classifier Probability that a Pixel is "Man-Made". Top: Complete probability overlay; Bottom: Overlay thresholded around 60% likelihood.

In addition to pre-processing images to address resolution, we investigated using Contrast Limited Adaptive Histogram Equalization (CLAHE) to address variable lighting conditions from the Kraken Solar Flare Mini underwater light. Applying CLAHE to the images before calculating predictions from the DCNN had noticeable but inconsistent effects from one dataset to the next. For the largest and most diverse image set (5-11), CLAHE appears to have reduced the number of false negatives - e.g. image patches with munitions assigned a low munition detection probability. In contrast, CLAHE appears to have substantially increased the number of false negatives for the 5-12 dataset. For the 5-13 image set, CLAHE resulted in only 3 (out of 100) image patches resulting in a munition detection probability greater than 80%, in contrast to the native intensity images yielding more than 20. While use of CLAHE had inconsistent effects on the predictions, it

did consistently make the images easier to interpret to the human eye, making it useful for visualization.

5-11 represented the largest and most diverse source of images. Example image crops organized by prediction from the DCNN are shown in Figure 25 and Figure 26. The DCNN returns a prediction confidence probability. Nominally, a probability of 90% would indicate that 9 out of 10 images contain a munition (likewise 50% would be half and half and 10% would be 1 out of 10). While DCNNs do not always precisely exhibit this behavior, they usually follow the general trend, and 50% confidence is often chosen as a threshold for detection. Despite this DCNN not being explicitly trained to detect munitions the trend is clearly present and based on a detection threshold of 50%, there are relatively few false positives as well as false negatives in this image set and they tend to be correlated with prediction confidence. This is also true with the 5-12 image set (no CLAHE) as seen in Figure 27. However, as seen in Figure 28, use of CLAHE has a significant negative effect on DCNN performance for 5-12. Overall performance on the 5-13 image set was not as good as seen in Figure 29 (no CLAHE) and Figure 30 (with CLAHE), however their remains some correlation with prediction confidence especially at higher confidence values (>90%).

Overall, the DCNN performance is quite good, especially given that it was trained for an entirely different application using images taken above water and largely on land.

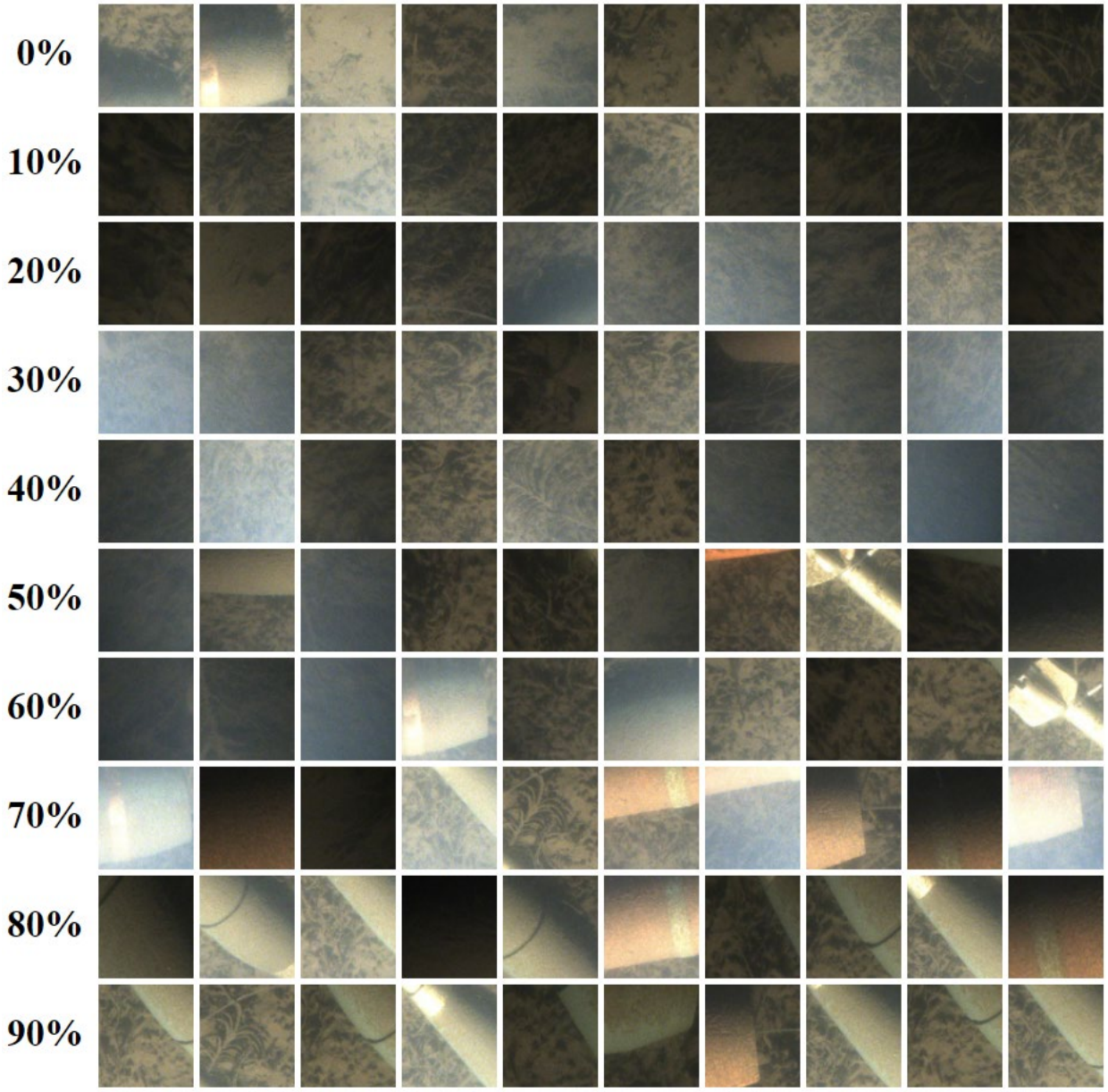


Figure 25. Shown Here are Randomly Selected 299x299 Image Patches Taken From the “5-11” SfM Image Set Organized by the DCNN Predicted Confidence Probability the Patch is of a Munition. Despite the DCNN not being explicitly trained to detect munitions, there is a clear correlation between the presence of munitions in the image patches and the DCNN confidence. That said, the image patch in the first row and second column is mislabeled despite image patches that would appear to a human to be very similar being assigned more appropriate confidences (see examples in the 60% and 70% groupings).

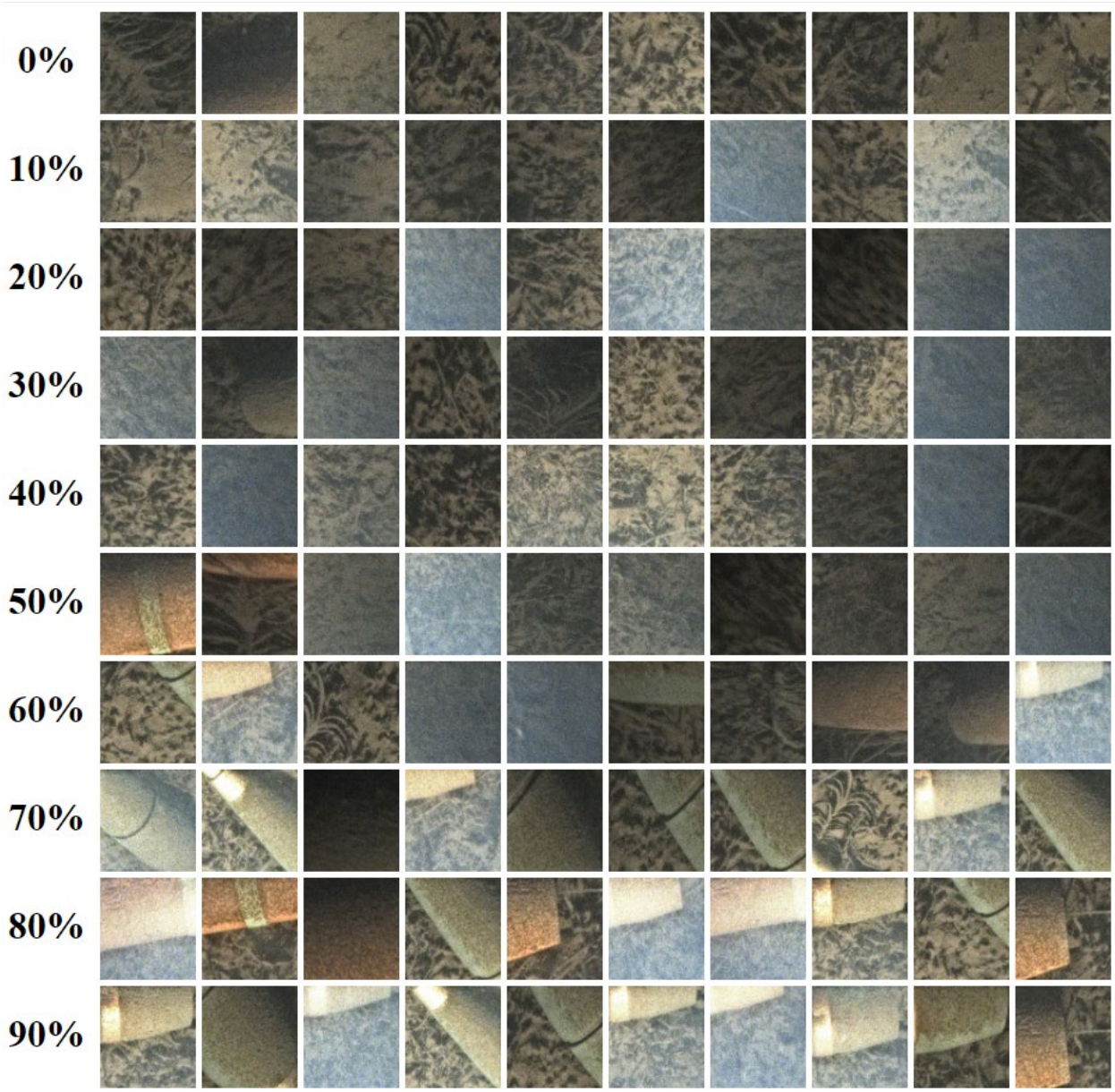


Figure 26. A Different Random Selection of the Same Source Image Patches as in Figure 25 are Shown Here Except That They Were Enhanced Using Contrast Limited Adaptive Histogram Equalization (CLAHE) Before Applying the DCNN. CLAHE improves the clarity to the human eye and for this image set also appears to have a slightly positive effect on the false positive and false negative rate when using 50% confidence as the detection threshold.

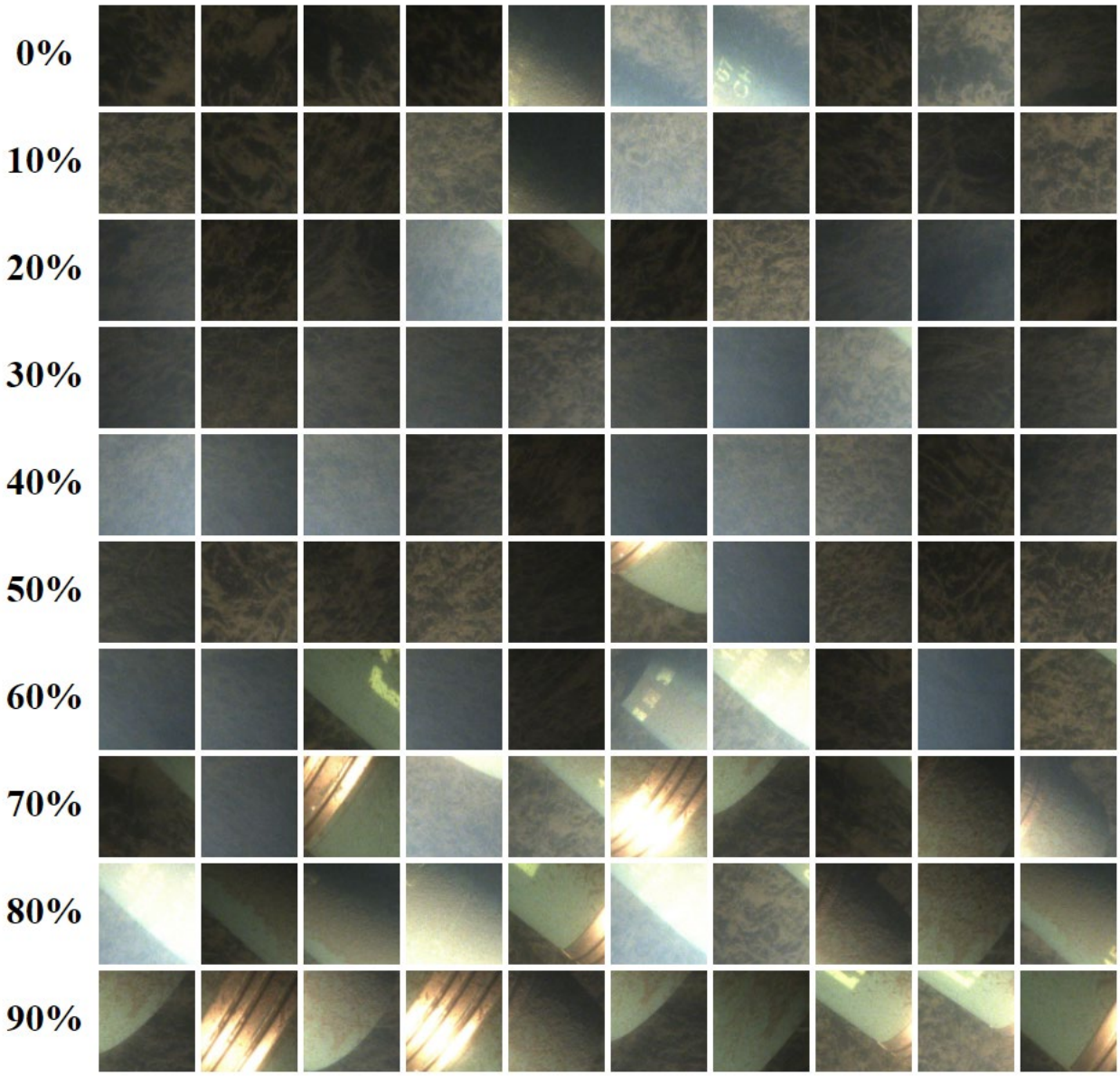


Figure 27. Random Selection of Image Patches From the “5-12” SfM Image Set, Organized by the DCNN Predicted Confidence Probability the Patch is of a Munition. As with the “5-11” image set, there is a clear correlation between the DCNN confidence and the presence of a munition in the image patch.

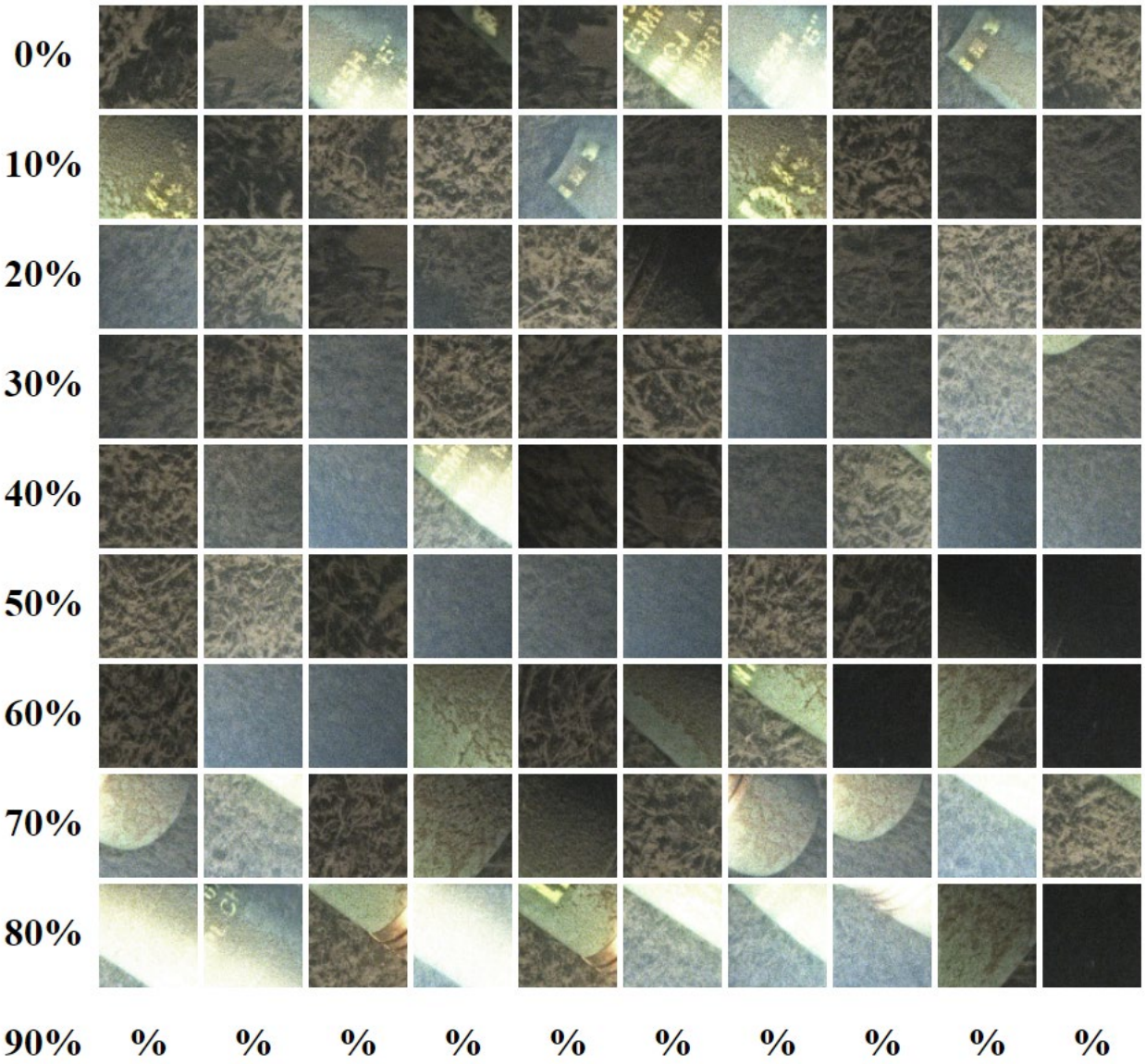


Figure 28. Random Selection of Image Patches From the “5-12” SfM Image Set That Have Been Pre-Processed Using CLAHE, Organized by the DCNN Predicted Confidence Probability the Patch is of a Munition. Use of CLAHE appears to have a negative effect on the DCNN performance in this case. In addition, none of the CLAHE pre-processed patches were predicted to have greater than 90% confidence. However, it still helps with visualization.

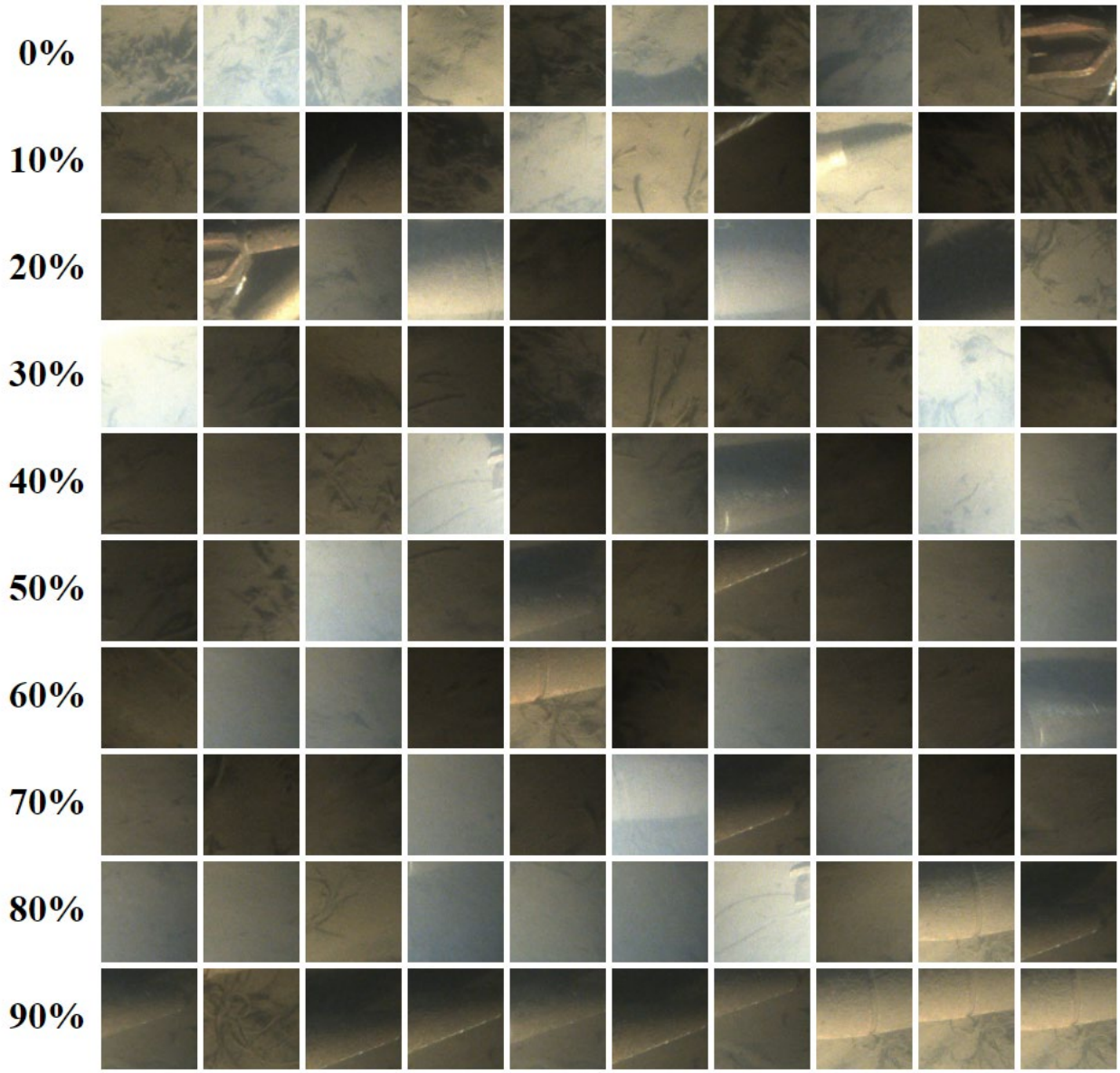


Figure 29. Random Selection of Image Patches From the “5-13” SfM Image Set, Organized by the DCNN Predicted Confidence Probability the Patch is of a Munition. With the exception of the patches showing 90% prediction confidence, the correlation between the DCNN confidence and the presence of a munition in the image patch is less apparent in this image set than 5-11 or 5-12. Interestingly, the DCNN appears to handle the coating of sand/dirt while failing to recognize fairly distinct features such as the bomb fins.

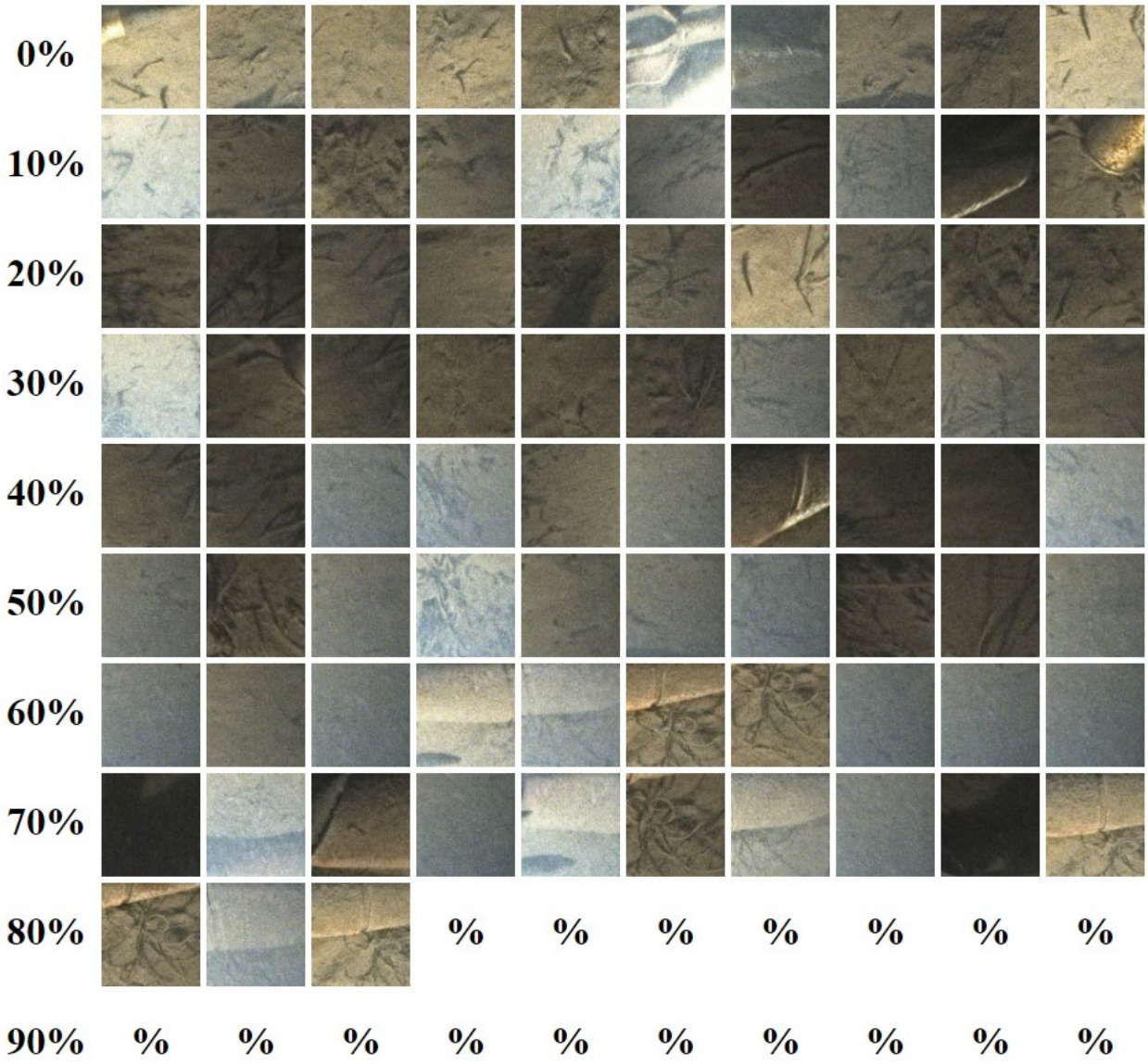


Figure 30. Random Selection of Image Patches From the "5-13" SfM Image Set That Have Been Pre-Processed Using CLAHE, Organized by the DCNN Predicted Confidence Probability the Patch is of a Munition. As with the "5-12" image set, CLAHE has resulted in no image patches receiving high confidence (>90%) from the DCNN prediction, and only 3 receiving confidence greater than 80%.

Potential Problems and Enhancements. While the overall performance of the DCNN is unexpectedly good, bearing in mind that it was trained for an entirely different purpose using images taken from an entirely different setting, there are some issues of note:

1. Similar image patches can result in different DCNN detection confidences.
2. Lighting adjustments (e.g. CLAHE) can have a significant effect on predictions.

Both issues can potentially be addressed by modifying the DCNN training procedure to include more training data from *Openimages* and by applying image data augmentation techniques. We would also re-select which image classes we use for training (e.g. there are lots of "man-made" objects in Figure 22's "Outdoors" category) Most of the *Openimages* classes have thousands of examples which is potentially an order of magnitude increase in the size of the training data originally used. Moreover, standard data augmentation techniques that apply random adjustments to the images (e.g. alterations in brightness, contrast, hue, scale as well as the addition of blur and noise) can be used to further enhance the diversity of the training data. Augmentation can also potentially be tailored to our end application where the light passing through the water can yield specific image effects. In addition, the convolution layer parameters can be included in optimization so that the network can be more specifically optimized to our application. While this can result in overfitting for networks trained with limited data, this should not be a problem with the expanded and augmented image set; and, given sufficient training data should yield more stable predictions.

In addition to improving the original DCNN training, if we specifically label images collected with our system as to the presence of munitions, these can be used to further refine the DCNN. Even a relatively small number of labeled images could increase performance. We envision an iterative procedure where our system is used with the existing DCNN. Each use yields more images of a target area. The existing DCNN is used to present an operator the most likely locations for the presence of munitions. Feedback supplied by the operator during normal use of the system is then used to automatically refine the DCNN and improve performance on subsequent runs (or reprocessing of the same run).

Summary. This is a promising approach to harness the power of neural network classifiers while minimizing the expensive and time-consuming collection of labeled training data. A follow-on effort should investigate this approach in more detail. In particular, the underlying classifier should be retrained with labeled data better aligned with "munition-like" and "not munition-like" (instead of our preexisting "weather" / "not weather" classifier used here). The user interface for reviewing OMD ATR hits should include a simple and easy procedure for the user to classify false positives and to review potential false negatives; this will enable on-the-fly retraining of the network for the specific site.

6.4.5 Automatic Target Detection – Machine Learning on Structured Light Data

We also developed a concept for a Machine Learning based approach to munition detection using structured light (SL) data collected by our system. The proposed approach has several advantages:

1. The primary advantage is that it has the potential to not require the arduous and expensive task, that often is required by supervised machine learning approaches, of human experts labeling data to create vast and diverse training data sets. Ideally some real data would be labeled to evaluate real world model performance, but this would be much more limited and could potentially be generated using the image based DCNN we have demonstrated.

2. Automated model refinement allows the system to improve over time with continued use by leveraging a growing collection of geographic site-specific SL scans.
3. SL is complimentary to the DCNN based image classification approach we have demonstrated. It provides detection capabilities based on 3D geometric information vs color-based detection. These two sources of detection can be used simultaneously for increased confidence/sensitivity. Alternatively, one can be used in lighting conditions that the other might struggle with.
4. Moreover, by jointly calibrating the cameras used for SfM and SL, data and predictions from the image based DCNN can be used to help improve those from an SL based model or to generate test data to evaluate model performance.
5. A modified version of the approach also has the potential to be applied to the coordinate/depth data from SfM and likewise combined with the image based DCNN classification to yield multiple predictions for each point in the SfM point cloud.

Machine Learning Approach. The basic approach is compatible with multiple deep learning model types and architectures, including DL models used for sequence data (1D DCNNs, LSTMs, Transformers) as well as 2D Deep Convolution Neural Networks. In addition, it can be used for classification or segmentation as illustrated in Figure 31 and Figure 32 respectively.

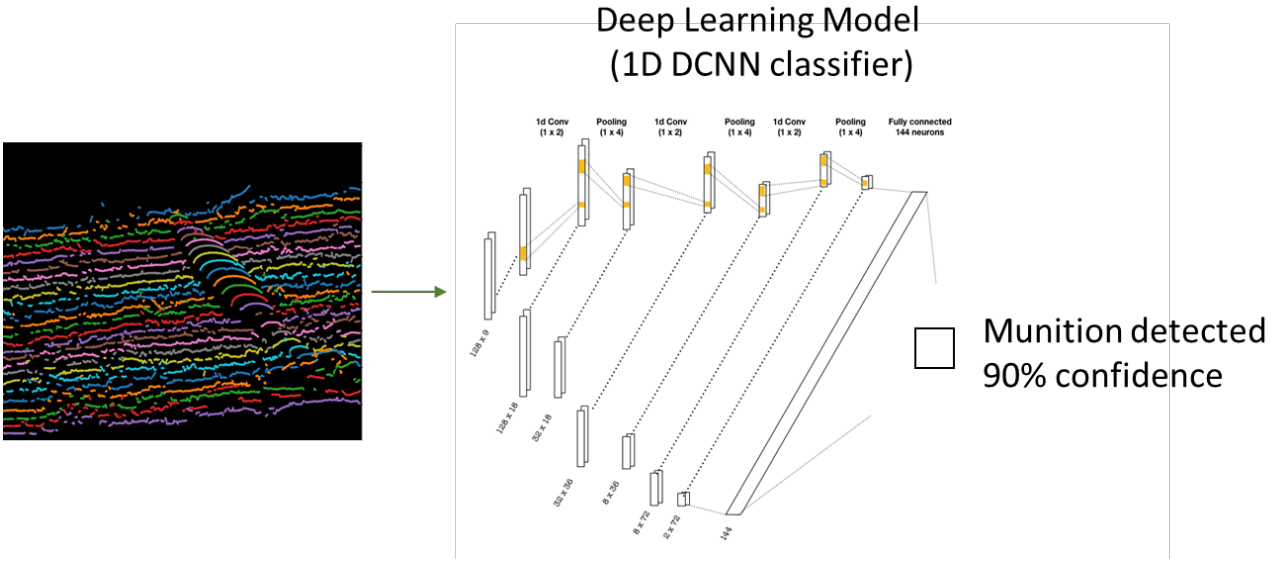


Figure 31. Example of a Deep Learning Classifier Used to Detect the Presence of Munitions in a patch of Laser Lines Produced from Multiple Adjacent Image Scans from the Structured Light Sensor System. Shown here is a multi-channel 1D Deep Convolution Neural Network (DCNN) where each individual laser line would be represented in its own channel as a sequence of depth values. Alternatively, a single channel 2D DCNN could be used. Output from such a model would be a single probability prediction that the input scene contains a munition.

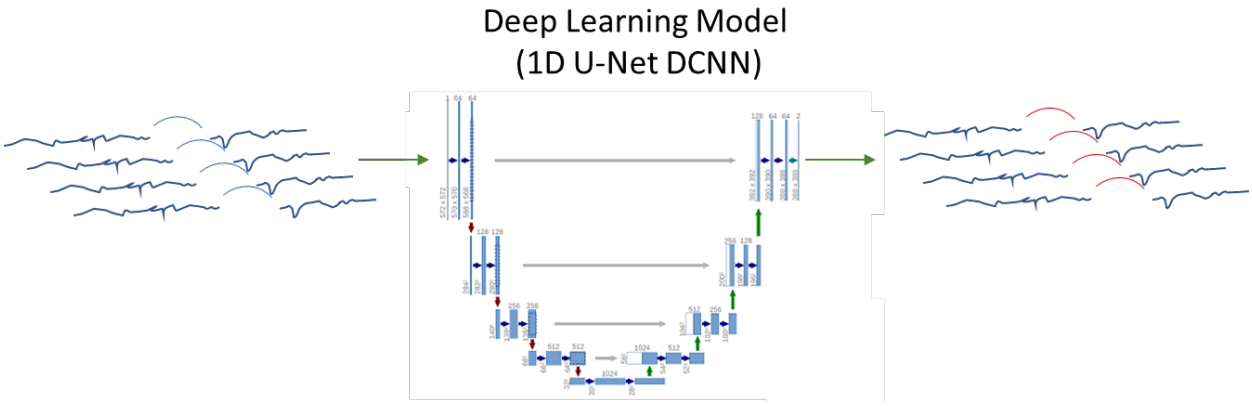


Figure 32. Example of a Deep Learning Segmentation Model. Like the classification model shown in Figure 31 the input would be a patch of adjacent laser lines. In this case, however, the output would be a probability for each point in the patch of laser lines rather than a single probability for the entire patch. Above, the red line segments in the output indicate the points that the model believes represent a munition based on a probability threshold. As with classification a multi-channel 1D model or single channel 2D model could be used. Shown above is a U-Net architecture segmentation DCNN originally developed for medical imagery but easily adapted to our problem.

Multi-Channel 1D Sequence. This approach treats each laser line scan as a 1D sequence of depth values. Multiple adjacent laser lines are fed into the ML detection algorithm so that more complicated higher dimensional geometric shapes can be modeled.

The proposed approach would use a multi-channel 1-Dimensional Deep Convolution Neural Network (1D-DCNN), however other sequence based Deep Learning models could potentially be employed such as those based on RNNs, LSMTs, or Transformers.

The problem could be formulated as classification or segmentation. Classification would yield a single confidence value that a munition is somewhere in the input patch of laser-lines. Segmentation would extend this to yield such confidence values for each point in the laser-line patch.

2D DCNN. Instead of using multi-channel 1D models, a single channel 2D DCNN (classification or segmentation) could be employed. 2D DCNNs are in fact what is typically used for image data and they have been shown to be extremely effective for these problems. In this case, multiple channels are typically used to represent the multiple colors in the images (e.g., RGB). Since we only have one channel of information (depth), we do not necessarily need a 2D DCNN which has more parameters and therefore more computational complexity which translates to more cost for inference and training.

In addition, each laser line is from a single image scan. While the images can be taken at a fixed frame rate, the distance between lines can vary due to fluctuations in speed of the collection system. Typically, 2D DCNNs are used on precisely gridded data where it is assumed the objects being detected are translation invariant across both dimensions. Treating the laser lines as separate channels could address any potential issues that result from this irregularity.

Synthetic Data for ML Training. For experienced practitioners, the most difficult aspect of supervised Machine Learning models is often the effort of having human experts label vast data collections sufficiently large enough to represent the diversity of data likely to be encountered when the model is used. To overcome this, computer generated synthetic data is sometimes employed to train a model. However, this only works if the synthetic data is sufficiently realistic. By combining actual data collected by our SL system of real sites with three dimensional models of munition targets of interest, and by employing data augmentation techniques, it may be possible to generate sufficiently realistic "synthetic" data to train a SL based classification or segmentation model.

The basic concept for our training data generation (depicted in Figure 33) is to draw on a library of SL scans collected at sites of interest depicting "background" objects and material (e.g. sand, vegetation, rocks, coral). Initially, there would be some effort to collect this data, however, normal operation of the system would continue to enhance this library. More importantly, the difficult and arduous task of labeling the data would not be needed. Instead, a library of three-dimensional representations of target munitions of interest would be used to insert synthetic targets into the background scans. Accomplishing this for the SL data is much easier than it would be for arbitrary generic 3D scenes with color because:

1. SL data only has a depth channel and no color channels. This drastically decreases the complexity of the target data, which effects model complexity and ability for the model to generalize. It also decreases the complexity of any random augmentation alterations we might employ to model variations in target data (e.g., covering of sand, dings, and other surface defects).

2. SL data represents a ground surface and not an arbitrary 3D point cloud. This makes it easier to place the synthetic targets in the real scene by using depth sampling of aligned geometries.

3. Background SL data can be collected from real sites of interest rather than needing to generate realistic synthetic background data for all possible scenes.

The process for training data generation leverages the precise 3D coordinates (in known absolute distance units) of the points in the laser lines of background scans from our calibrated SL system so that the coordinate systems of the scans and the 3D munition models can be realistically aligned. A single example of training data can then be generated with the following process which is depicted in Figure 33:

1. Select a random patch of laser lines just sufficiently large enough to contain most if not all munitions found in the library of target munition models. In the case of a classification deep learning model this patch would then either be used as an example input representing no munition targets present or the following steps would be used to alter it to represent a positive example for training.
2. Randomly select a single target munition from the library of 3D models.

3. Generate a random rotation/translation transform between the coordinate system of the laser lines and the target munition 3D model subject to certain constraints such as sufficient overlap of the munition and the laser line patch. A constrained random scaling transform could also be employed to increase sensitivity to munitions of different sizes. This step can be viewed as an example of data augmentation.
4. Use the known x,y coordinates of each depth value in the patch of laser lines to sample the aligned 3D model. In the case that no point of the munition target exists for the x,y coordinate, the background laser line depth value is used. In the case of one or more points in the munition target existing for a single x,y point, the smallest depth value from the munition model and background laser line scan would be used so as to represent the point observable to the SL scanner (i.e. the closest). Concurrently, the decision to use a depth value from the background laser line patch vs. the target munition would be noted and used as labeled training data for segmentation.
5. The depth values from the munition target can then be randomly modified to model a coating of sand or possibly a surface defect.

DCNN training would then involve generating batches of these random training data examples to optimize the model parameters. Ideally, real data, including examples of SL scans of actual munitions, would be used as validation data during training and as test data to evaluate the realistic performance of the deep learning model.

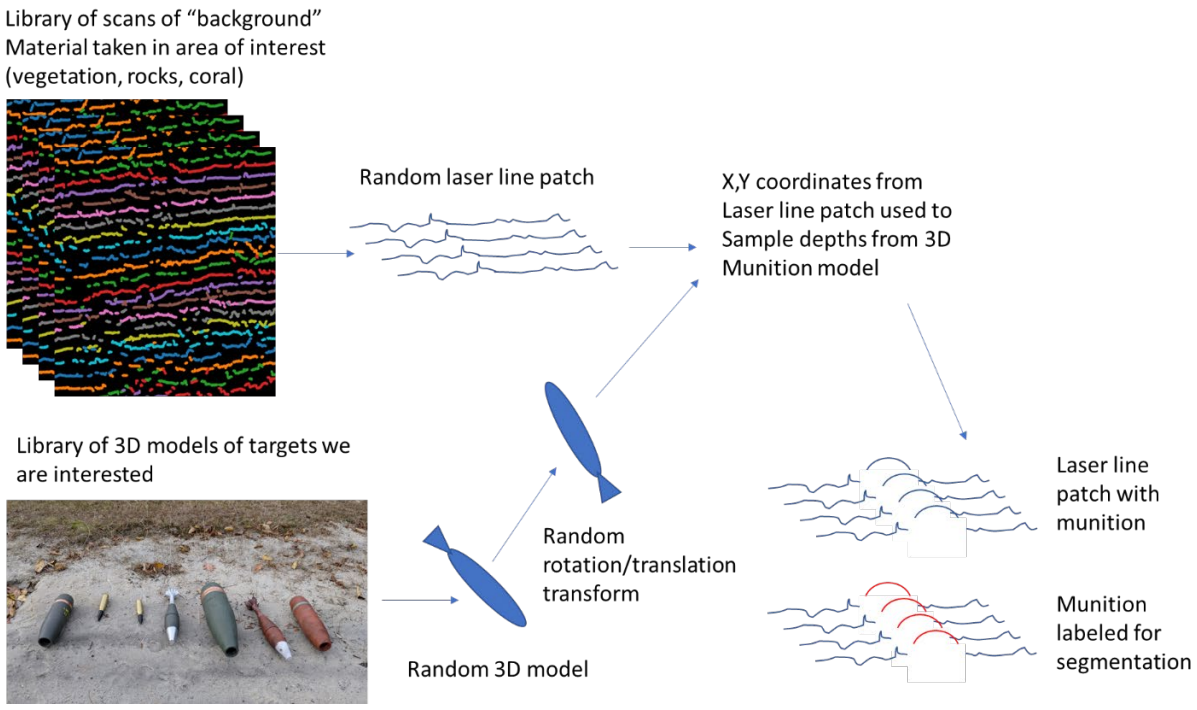


Figure 33. The diagram above illustrates the process of combining real SL scans of background material and objects (e.g. sand, vegetation, rocks, and coral) with a library of 3D models of munition targets in order to generate a realistic synthetic data example that can be used to train an SL based Deep Learning segmentation model. A similar and simpler procedure can be used to generate training data for classification.

Potential Problems and Enhancements. There are two potential problems we anticipate with the SL system:

1. The SL data points can be somewhat irregularly spaced. This can be due to 1) a change in velocity of the collection system between SL scans, 2) irregular spacing of pixels within a single scan due to fanning from the camera to the ocean floor.
2. The synthetic data could be insufficiently realistic.

Deep Learning models, particularly DCNNs, are usually used on regularly structured and evenly spaced data such as images that are represented as a grid of evenly spaced pixel values, or time series data with a constant sampling rate. That said, the irregularity in the pixel spacing within a single laser line scan would be similar to what would be observed in a normal RGB image and therefore should only be a problem if our synthetic data does not behave realistically in this regard. However, precise calibration of our SL system should allow us to directly model this by providing real world coordinate data of each pixel in the laser line scan that we can use to realistically place the munition target models. The more likely problem comes from uneven spacing between scans due to velocity changes.

We believe using multi-channel 1D convolutions (or other multi-channel sequence DL algorithms) over a 2D DCNN has a better potential for addressing this inconsistency because the convolutions are 1D rather than 2D and therefore due not make any assumptions about the regularity in the dimension representing consecutive scans.

As an alternative, a technique like CoordConv[6] can be used to explicitly model coordinates. In this case coordinate data (relative to the SL collection system) would be supplied to the DCNNs (1D or 2D) as separate data channels which would allow the DCNNs to parametrically model any irregularity effects directly.

Some deficiency in the realism of the synthetic data can be overcome by following up training on the synthetic data with fine-tuning of the model on a smaller collection of real data. The synthetic data ensures the model can handle the diversity of data that could be encountered during use of the model, and the real data ensures the model can handle the smaller scale details found in real data as well as any consistent characteristics of the real data not present in the synthetic data. For-example, while our synthetic data would allow the model to detect the geometry of the munition targets it would not explicitly include effects such as dirt/sand building up around the munition. This should not matter for detecting the bulk of the munition but may affect how precisely the munition is segmented at the border between the munition and the background dirt/sand. Continued refinement with real labeled data would enhance the detection capabilities so as to capture effects like this.

Summary. This approach is promising and worthy of investigation in a follow-on effort (or, resource-permitting, during the time between our Draft and Final Reports). With some additional

[6] https://arxiv.org/abs/1807.03247

pre-processing steps this approach may also be applicable to the structure from motion topology (depth) data.

7 CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH

7.1 CONCLUSIONS

We have been able to draw a number of conclusions and lessons learned from our work. In particular:

- Our general imaging and reconstruction approaches are feasible and well suited for exposed munitions in shallow-water environments. Although the data set is, as anticipated, limited, munitions are clearly visible in the raw color images, composite structure-from-motion point clouds, and structured-light point clouds.
- The illumination of the structure-from-motion (color camera) scene needs improvement. There are strong lighting gradients, strong shadowing, and particle-backscatter that confound the SfM reconstruction and DCNN munition detection algorithms. Dr. Jaffe's recommendations for improving the lighting (and image quality in general) should be implemented in the next-generation system.
- The signal to noise ratio of the structured light imaging system was, as expected given the hardware we used, poor. Increasing the laser power and the addition of a band-pass filter will vastly improve this. This has low technical risk.
- A number of munition detection algorithms have promise. We have identified shape-based approaches for the SfM point clouds; texture-based approaches for the raw images; machine learning (DCNNs) for the raw camera images; and DCNNs for the structured light data.
- Of these, the DCNN algorithms are likely the most generalizable. Training neural networks is typically the most time-consuming and expensive part of the development process. As we have shown here, applying networks trained on existing data sets (in-air, no munitions) can serve as a very reasonable starting point if the network outputs are appropriately categorized. These can be retrained on-the-fly with human-in-loop review of the data.
- We also now appreciate (even more) the need for an intuitive and useful user interface for displaying survey results (and for possible on-the-fly retraining of munition detection algorithms). We will draft user interface concepts for our (to be scheduled) Final Results presentation.

7.2 IMPLICATIONS FOR FUTURE RESEARCH

7.2.1 Potential Concept of Operations

Over the course of this effort, we have refined our vision for the OMD's concept of operations. Participating in the December 2020 SERDP Symposium was particularly valuable for drafting this vision.

In the realm of all munition-contaminated underwater sites, there is particular concern for shallow water sites where the munitions are exposed: the public is most likely to interact with these munitions. Indeed, some remediation action plans are intentionally limited to exposed munitions.

Currently, a typical remedial investigation approach involves: 1) locating targets of interest (TOI) via acoustic (or, sometimes, electromagnetic) approaches followed by 2) diver investigation of these TOI. Diver investigation is expensive, time-consuming, and risky. Furthermore, georeferencing diver-collected data is challenging.

The OMD offers a single alternative to both steps. The OMD, deployed from an unmanned surface vehicle (USV) or unmanned/autonomous underwater vehicle (UUV / AUV) can survey the site, detect TOI, and rank their probability of being munitions. OMD data (photographs and bathymetric topologies) is human readable, enabling a simple post-processing interface for quickly and intuitively verifying the TOI classification.

As the OMD is designed to operate at relatively close standoff distances (a few meters), survey speeds (area coverage rate) will likely be slower than acoustic approaches. As such, an alternative concept of operations is for the OMD to follow a traditional sonar survey. The OMD will selectively scan TOI coordinates flagged by the sonar survey. This approach has the added benefit of linking sonar and OMD data which will, over time, build a labeled sonar "image" data set – useful for training a future classifier of sonar data.

7.2.2 Recommendations for Future Research and Development

The OMD remains well suited to close operational gaps with underwater munition detection and classification. It is particularly well suited for shallow water environments (that are traditionally challenging for acoustic imaging approaches) and for exposed munitions (which are most at risk of interaction with the public). We recommend a follow-on effort to develop a fully functional prototype. This prototype should feature:

- Both structure from motion and structure light imaging. Structure from motion (and the raw images the feed the SfM algorithms) enables a broader range of automatic target recognition technologies. However, structured light imaging will work over a wider range of water conditions (turbidity), a particular concern for shallow waters.

- Improved image quality. Implementing the imaging recommendations identified on this effort should improve both raw image quality and reconstructed three-dimensional point clouds.

- Integration with an unmanned host vehicle. While the basic OMD technology is agonistic to the specific host vehicle (surface or underwater), an unmanned surface vehicle is a good candidate for the next prototype. Surface vehicles tend to be less expensive, have more flexible payload interfaces, and offer straightforward geolocation and route finding via GPS. Surface vehicles are also well suited for a number of potential test beds. ROVs, AUVs, or towed platforms would allow the OMD to be used in deeper water; no design decisions should be made that exclude eventual use on these alternate platforms.

- Automatic Target Recognition. We identified a number ATR technologies in this effort worthy of future investigation. In particular, curvature (geometry) based mapping and deep convolution neural network classification are likely the most generalizable and the most promising. A follow-on effort should continue to develop these approaches for this application.

- Simple and Intuitive User Interface. OMD data should be displayed on a user interface that features georeferencing, ranking of TOI by probability of being "munition-like", and access to the underlying images and topology for verification or rejection of the classification.

The prototype should then be validated in the open water:

- Local Testing. This testing will similar to the testing done on the initial effort. We recommend adding additional targets, conditioning the target surfaces to be more representative of years-old UXO, and adding clutter targets (e.g. tires, scuba tanks, etc.).

- Dedicated UXO Test Site. SERDP is developing a number of underwater UXO test beds. In particular, the Hawaii Munitions Test Range Complex is particularly well suited for the OMD. It is a shallow water site with good water clarity, significant natural clutter, and can be readily scanned from a surface vehicle. Alternatively, sites with real and documented UXO contamination (such as Joe English Pond in New Boston, NH or the Island of Vieques, Puerto Rico).

- Other Test Sites. If necessary, additional work could include the addition of more challenging test sites (those with more turbid water or those deep enough to require integration with an undersea vehicle).

This plan should be sufficient to demonstrate the utility of the OMD as a tool for underwater munition site remediation.

8 LITERATURE CITED

Davies-Colley, R. J., "Optical properties and reflectance spectra of 3 shallow lakes obtained from a spectrophotometric study," *New Zealand Journal of Marine and Freshwater Research*, 1983, Dec, Vol. 17, No. 4, pp. 445-59.

Hatcher, G. A., Warrick, J. A., Ritchie, A. C., Dailey, E. T., Zawada, D. G., Kranenburg, C., Yates, K. K., "Accurate Bathymetric Maps from Underwater Digital Imagery Without Ground Control," *Frontiers in Marine Science*, 2020 Jun, Vol. 26 No. 7, pp. 525.

Jaffe, J. S., "Computer modeling and the design of optimal underwater imaging systems". *IEEE Journal of Oceanic Engineering*, 1990, Apr, Vol. 15, No. 2, pp. 101-11.

Jaffe, J. S., "Monte Carlo modeling of underwater-image formation: validity of the linear and small-angle approximations," *Applied Optics*, Vol. 34, No. 24, 1995, pp. 5413-5421.

Jaffe J. S., Moore K. D, McLean J., Strand M. P., "Underwater optical imaging: status and prospects," *Oceanography*, 2001, Jan, Vol. 14, No. 3, pp. 66-76.

Jaffe, J. S., "Underwater optical imaging: the past, the present, and the prospects," *IEEE Journal of Oceanic Engineering*, 2014, Oct, Vol. 40, No. 3, pp. 683-700.

Jaffe, J. S., "Performance bounds on synchronous laser line scan systems" *Optics express*, 2005, Vol. 13, No. 3, pp. 738-748.

Jerlov, N. G., "Marine Optics," *Elsevier*, 1976.

Kirk, J. T., "Yellow substance (gelbstoff) and its contribution to the attenuation of photosynthetically active radiation in some inland and coastal south-eastern Australian waters," *Marine and Freshwater Research*, 1976, Vol. 27, No. 1, pp. 61-71.

Mobley, C.S., (OCEAN OPTICS Web Book) (https://www.oceanopticsbook.info), 2021.

Strategic Environmental Research and Development Program, "SERDP workshop on acoustic detection and classification of munitions in the underwater environment," SERDP-ESTCP Final Report, 2018.

van Scheltinga, R. C., et al., "Observations of dune interactions from DEMs using through-water Structure from Motion," *Geomorphology*, Vol. 359, 2020.

Zege, Ė. P., Ivanov, A.P., Katsev, I. L., "Image transfer through a scattering medium," *Springer Verlag*, 1991.

APPENDICIES

9 UNDERWATER IMAGING PRIMER (PREPARED BY OUR CONSULTANT, JULES JAFFE)

9.1 CHARACTERIZATION OF OPTICAL PROPERTIES OF THE WATER COLUMN THAT AFFECT UNDERWATER IMAGING.

9.1.1 Introduction

As has been known for some time, starting with the excellent work that was accomplished at the Scripps Inst. of Oceanography's Visibility Lab in the 1960's the important parameters for optical characterization of the water column that will predict imaging are the scalar properties of total attenuation, total absorption, and total scatter. These are accompanied by the vector function that is the Volume Scatter Function, or VSF that is essentially a probability density function (per inverse meter) for a light beam to be scattered in a given direction from zero degrees for forward scatter to 180 degrees for backscatter. Together, these inherent optical properties can then be used to parameterize the moderately complex distribution of light from either natural or artificial sources as well as the point spread function (PSF) that will characterize the loss of image resolution that is due to the small angle scatter of light, a result of the distribution of particles that have an index of refraction different of that from the medium that they are suspended in.

As a function of the Volume Scatter Function, an important characteristic of light propagation is the amount of both spreading of the beam in the forward direction as well as the backscatter of light that produces a veiling glare for observing images. The former is most important for the laser light sheet, while the latter effect will reduce the contrast for the SfM camera system. The environments that are being considered for the OMD range from open sea water to coastal water to lakes.

Given that there are several references in on-line web sites such as the OCEAN OPTICS Web Book (Mobley, 2021) and the authors own articles (Jaffe, J. S., 1990, 2001, 2005, 2014) there seems to be little need to go through the physics of light propagation and its consequences for underwater optical imaging in environments that range from highly turbid to exceptionally clear. Instead, here we will summarize the consequences for imaging variability in the context of the specific goals of this program: That is, to use laser sheet illumination for estimation of the bottom topography as well as visibility and the 3-dimensional information that can be used to achieve the Structure From Motion (SfM) goals prescribed.

9.1.2 Optimal Choice of Wavelength as Dependent on the Environment (Marine)

An important aspect of underwater imaging, whether in lakes or in a marine environment, is the choice of optimal wavelength of the light. As the OMD is envisioned to operate in both environments, the system's flexibility in choosing the incident light for both laser sheet illumination and SfM is important. Considering first, marine environments, Table 2 (extracted Jerlov (1976)) contains a brief summary of the three oceanic environments (transparent, least transparent, and average transparency) with the fraction of light that is lost in 1 meter as a function of wavelength. As illustrated, in clear water, the least light is lost at the blue end of the spectrum.

This would occur in tropical areas that are typically low in nutrients such as offshore Hawaii. Here, total attenuation lengths (the range at which e^{-1} of the light is lost) can be as high as 20 meters. This would then favor the use of blue light at wavelengths of (420 – 460 nm) although using light at slightly longer wavelengths, in the blue green, would only result in slightly less light loss. We note that in coastal areas, that are the most relevant to the operational environment of the OMD, the transparency window shifts more to the green part of the spectrum, that makes the use of blue green lasers ideal.

Table 2. Loss of light at 1 meter as a function of wavelength.

Loss of light (percent) in one metre of seawater*								
	violet		blue-green			yellow	orange	red
wavelength (micrometre)	0.30	0.40	0.46	0.50	0.54	0.58	0.64	0.70
oceanic water, most transparent	16%	4%	2%	3%	5%	9%	29%	42%
oceanic water, least transparent	57%	16%	11%	10%	13%	19%	36%	55%
coastal water, average		63%	37%	29%	28%	30%	45%	74%
*According to Jerlov.								

9.1.3 Optimal Choice of Wavelength as Dependent on the Environment (Lakes)

The environmental situation in lakes is quite different that in marine environments. A brief survey reveals that although there are plankton blooms, lakes are subject to a yellowish substance called gelbstoff (Kirk, 1976), that is a component of dissolved organic matter that occurs from tannin released from decaying detritus. This will cause the water to appear green, yellow green, and then brown. Figure 34, taken from Davies-Colley (1983) shows that the absorbance (ratio of logarithm of incident light to transmitted light) of lakes in New Zealand, as described using absorbance, decreases toward the red end of the spectrum. This presents a very different picture for imaging in lakes that should be taken into consideration in determining the optimal wavelength to use for the OMD. I would clearly favor using light that is more towards red or orange. The volume scatter functions and those wavelengths are also more favorable as the ration of wavelength to size of small particles implies more forward and less side and back scatter.

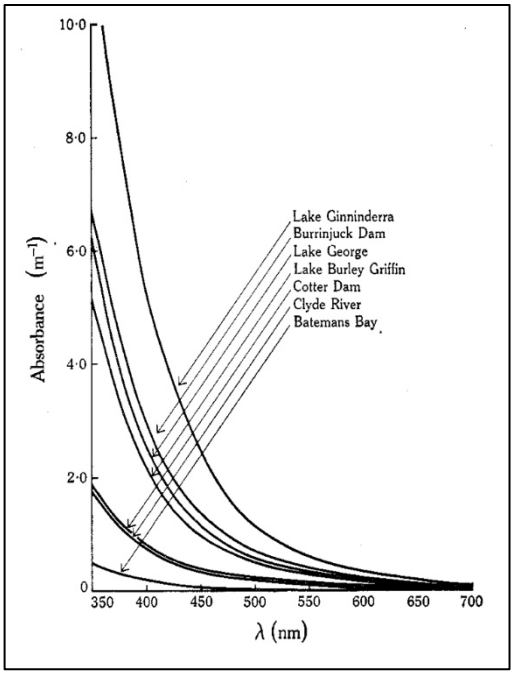


Figure 34. Absorbance spectra in New Zealand lakes. Davies-Colley (1983)

10 GEOMETRIC CONSIDERATIONS IN UNDERWATER IMAGING

10.1 LASER LIGHT STRIPE IMAGING

In considering the potential performance of a system that is based on the propagation of a laser sheet that is then imaged by an underwater camera system, we note that the system can provide both reflectance information along the reflection of the light stripe in addition to the use of triangulation in order to compute range, or bottom topography. The basic geometry of the situation and its incorporation into a light stripe system that was deployed in the Gulf of Mexico to consider changes in sea floor micro-bathymetry (Moore and Jaffe, 2002) is depicted in Figure 35.

System geometry was a bit complicated as a 1024 CCD chip was used to estimate the location of the laser beam in angle that was reflected from the bottom and, hence, the depth of the system from a large armature that permitted repeated scanning. A relatively powerful blue green (532 nm) 3W laser was used that possessed minimal beam divergence that is essential for illuminating a small spot on the sea floor. Calibration measurements for the system showed that the absolute depth resolution is less than about 5 mm at 3 m, about 10 mm at 5 m, and about 31 mm at 10 m range. Averaging of successive bathymetric profiles then produced a resolution of less than 1 mm. Approximate separation of source and receiver were 45 cm and the beam divergence of the laser was extremely small, on the order of milliradians. Together with the narrow beam illumination beam, and the given camera light separation, the near field backscatter was mostly eliminated. Hence, the resultant images are constrained by the point spread function of the laser. Having a small divergence angle and given the close range, the effective point spread function of the system was extremely small at this close range. In

addition, the centroid of the reflected energy or irradiance, is the important quantity that is measured. This permits localization to the 3 dB points of the imaged beam, yielding the superior resolution, as described. Parenthetically, we note that the system, over the course of 9 days, recorded that the sea floor ripples that were the result of a previous storm were reduced from a 2.5 cm amplitude to a 2.0 cm amplitude. In the absence of strong currents, as evidenced by the ADCP, the decrease in bottom micro-bathymetic wavelength was the result of the biological resuspension of sediments by marine organisms.

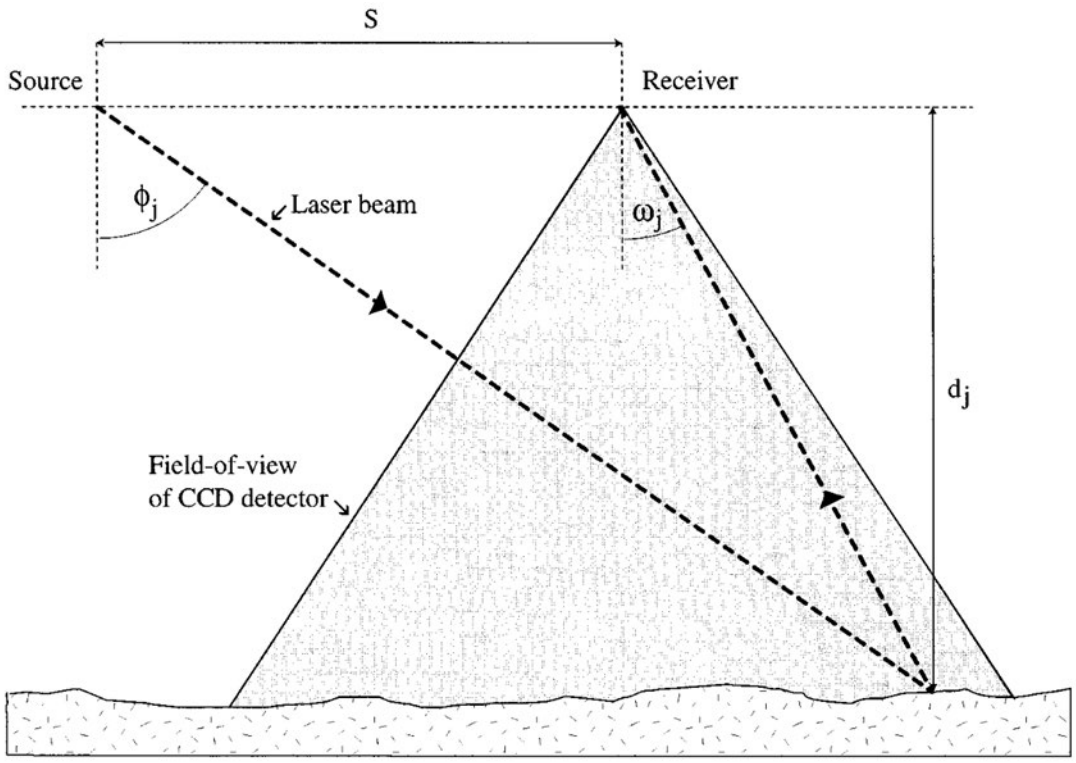


Figure 35. The geometry of the 3D Sea Scan Imaging System that mimics a laser light sheet illumination (Moore and Jaffe, 2007)

In making a choice to use a laser sheet for structured illumination or a scanning laser that has a point like beam, we note that the 532 nm wavelength laser wavelength is ideal of coastal waters as there is ample data that indicates that there is minimal absorption at this wavelength. In addition, given the current generation of laser diodes, the divergence angle of the beam can be very narrow (.5 – 1 mrad), that will confer a reduction of backscatter via separation of source and receiver. Using a laser sheet, the point spread function, or impulse response, will be wider in the cross-track dimension than in the along track dimension, yielding a decrease in lateral image resolution. However, the scan rate of such systems is one line per frame. This precludes the need for scanning at each and every point and aids in the simplicity of the system design. A more complete analysis of underwater optical imaging systems and their resultant point spread functions in covered in Jaffe (2005). One thing that works in favor of the laser sheet, in this application, is that at the close range, the point spread function would likely be small enough to define good range resolution as the beam thickness is, to first order spread linearly (Jaffe, 1995). We also note, most

importantly, that the backscatter from the laser beam will be imaged, for the most part, separate from the interesting area where the beam is incident on the sea floor.

10.2 Flood Illumination for Determining both Reflectance and Structure from Motion

In the other mode of imaging, that is for Structure from Motion, a broad beam of light is projected that permits illumination of a larger part of the scene. As is well known, broad illumination in turbid waters most often results in a large degree of backscatter that is considered to be a "veiling glow" that leads to a decrease of image contrast. The closest example that is also well known is that when driving a car in fog, turning one's bright lights on reduces contrast. In the context of evaluating the success of the CREARE effort and future applications of SfM, it is critical to examine the placement of cameras and lights to optimize the system's imaging performance. Jaffe's own UNCLES program that was developed the in the mid 1980's is a computer simulation that permits a variety of imaging options to be run with simulated results that contain signal to noise and contrast. Here, we note that the OMD system is envisioned to be at rather close ranges of meters. This helps to increase incident light that is not absorbed and also, permits the use of wider beam patterns via placement of the camera and lights at distances, on this relatively small platform.