



Challenges of Labelling Unknown Seabed Munition Dumpsites from Acoustic and Optical Surveys: A Case Study at Skagerrak

Oscar Bryan ^{1,*}, Roy Edgar Hansen ^{2,3}, Tom S. F. Haines ¹ , Narada Warakagoda ^{2,3} and Alan Hunter ^{1,3}

¹ Faculty of Engineering and Design, University of Bath, Bath BA2 7AY, UK; tsfh20@bath.ac.uk (T.S.F.H.); ajh210@bath.ac.uk (A.H.)

² Norwegian Defence Research Establishment (FFI), 2007 Kjeller, Norway; rhn@ifi.uio.no (R.E.H.); narada-dilp.warakagoda@ffi.no (N.W.)

³ Faculty of Mathematics and Natural Sciences, University of Oslo, 0315 Oslo, Norway

* Correspondence: ob300@bath.ac.uk

Abstract: The disposal of unexploded ordnance (UXOs) at sea is a global problem. The mapping and remediation of historic UXOs can be assisted by autonomous underwater vehicles (AUVs) carrying sensor payloads such as synthetic aperture sonar (SAS) and optical cameras. AUVs can image large areas of the seafloor in high resolution, motivating an automated approach to UXO detection. Modern methods commonly use supervised machine learning which requires labelled examples from which to learn. This work investigates the often-overlooked labelling process and resulting dataset using an example historic UXO dumpsite at Skagerrak. A counterintuitive finding of this work is that optical images cannot be relied on for ground truth as a significant number of UXOs visible in SAS images are not in optical images, presumed buried. Given the lack of ground truth, we use an ordinal labelling scheme to incorporate a measure of labeller uncertainty. We validate this labelling regime by quantifying label accuracy compared to optical labels with high confidence. Using this approach, we explore different taxonomies and conclude that grouping objects into shells, bombs, debris, and natural gave the best trade-off between accuracy and discrimination.

Keywords: synthetic aperture sonar (SAS); unexplored ordnance (UXO); machine learning



Citation: Bryan, O.; Hansen, R.E.; Haines, T.S.F.; Warakagoda, N.; Hunter, A. Challenges of Labelling Unknown Seabed Munition Dumpsites from Acoustic and Optical Surveys: A Case Study at Skagerrak. *Remote Sens.* **2022**, *14*, 2619. <https://doi.org/10.3390/rs14112619>

Academic Editors: Gil Rito Gonçalves and Umberto Andriolo

Received: 28 March 2022

Accepted: 5 May 2022

Published: 31 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the aftermath of World War II, it is estimated that over 700,000 tons of chemical weapons were dumped in European waters [1]. These historic UXOs are part of a worldwide problem caused by live firing and the disposal of surplus stocks. One such dumpsite used by British and American forces was Skagerrak. This dumpsite is located off the south coast of Norway at a depth of approximately 700 m. Here, the seabed is generally flat, consisting of fine sediment with sparse rock coverage. Visible craters from the sinking of ships laden with munitions suggest minimal sediment transport in the area [2]. On sinking, some ships stayed intact whilst others broke apart, shedding chemical munitions over a large area of seafloor. The large areas and inaccessible sites make locating and classifying UXOs on the seafloor challenging.

Despite the challenges, autonomous underwater vehicles are capable of surveying large areas of the seafloor using a range of sensing equipment: synthetic aperture sonar (SAS) [3], from multiple views combined with optical images, provides a rich source of information. The vast amount of data collected from these surveys necessitates automated approaches to target recognition (ATR). State-of-the-art approaches to ATR with imaging sonars tend to use deep learning trained in a supervised fashion [4–6]. This approach requires a large dataset of labelled examples for tuning the parameters of the deep learning model.

Safeguarding marine activities motivates understanding the location and extent of UXOs dumped at Skagerrak and elsewhere. Although a lot of studies focus on model

design and training regimes, the challenges of applying automated target recognition to historic UXOs are more fundamental. Placing known targets on the seafloor can be used for data collection. However, long exposure to the subsea environment and the impact of debris and piled objects limit the realism of placed targets. Additionally, identifying targets with divers [7] or recovering samples [8] from Skagerrak is difficult due to unknown amounts of chemical weapon agents remaining and the inaccessibility of the site. Therefore, no true ground truth exists, which is problematic for training and validating a supervised learning method. The labelling procedure and analysis of these labels is therefore an important consideration for guiding future machine learning efforts.

In this work, we label a historic UXO dumpsite from which we present and quantify considerations necessary for future machine learning research efforts. To this end, we develop a labelling methodology suitable for capturing label uncertainty given degraded targets and limited image quality. We aim to use the rich source of multi-view and multi-mode data available from Skagerrak to quantify the effect of look angle and sensor mode on target visibility. We also aim to quantify the human labelling effort and accuracy of those labels compared to high-quality optical images. The objective of this work is to provide a guide for future machine learning efforts and to inform the feasibility of mapping UXOs in terms of object visibility and taxonomy.

2. Autonomous Surveys

In 2015 and 2016, surveys covering 450 km² of seafloor were conducted at Skagerrak by the Norwegian Defence Research Establishment (FFI) [9]. This large survey was made possible by the use of a Kongsberg HUGIN autonomous underwater vehicle (AUV) fitted with an optical camera and synthetic aperture sonar; details of these sensors are in the following paragraphs. A total of 36 shipwrecks believed to contain chemical munitions were found and documented [9]. UXO types, along with examples of debris and natural clutter, imaged by SAS are shown in Figure 1. In 2019, a smaller subsection of the survey area was re-surveyed with an optical camera carried by the HUGIN AUV.

The optical sensor used was a Tilefish downward-looking camera with a 400 W strobe LED array [9]. The 10.7 megapixel camera captures small details, imaging the seafloor as a perspective projection of the 3D environment at 1 mm resolution [10]. However, the coverage rate is limited by a downward-looking geometry with a short stand-off, a result of the strong attenuation of electromagnetic waves in water (shown in Figure 2). For a typical AUV altitude of 6 m and speed of 2 m/s, the area coverage rate becomes 12 m²/s, or 0.043 km²/h [9]. At depth, a colocated illumination source is necessary; this arrangement does not produce a significant shadow when looking at proud objects. Additionally, sediment deposited on sunken objects produces reflectivity similar to that of the seafloor, leading to low-contrast images.

The other primary sensor used on the HUGIN vehicle was a HISAS synthetic aperture sonar with a centre frequency of 100 kHz and a 30 kHz bandwidth [11]. Active sonar emits an acoustic signal and infers information about the environment from the time delay and intensity of the echo [3]. Sonar resolution is limited by its aperture size; to improve resolution synthetic aperture sonar (SAS) is a technique for generating a large aperture on a moving vehicle. In this study, the FOCUS software developed by FFI was used to process the raw data (see [12] for more details on the SAS processing pipeline for the HUGIN AUV). The output typically used for target recognition, and indeed used in this study, is an intensity image where pixel values correspond to echo strength. It is an orthographic projection of the 3D space which has an aerial-photograph-like quality [3].

Good acoustic propagation in water permits a large effective range of 200 m per side (at 2 m/s vehicle speed) and hence a greater instantaneous coverage rate: 2.88 km²/h at 4 cm resolution [9]. The divergence of acoustic signal means that objects at a greater distance are ensonified by more pings. The synthetic array can therefore be longer, allowing range-independent resolution. SAS is typically operated at a shallow grazing angle, resulting in a strong reflection and shadow, highlighting proud objects (Figure 2). Furthermore, the

high density of metallic objects (e.g., UXOs) returns a strong acoustic echo, giving a good contrast to background sediment. The major drawback is limited resolution; this means small targets have poor detail, being represented by only a few pixels.

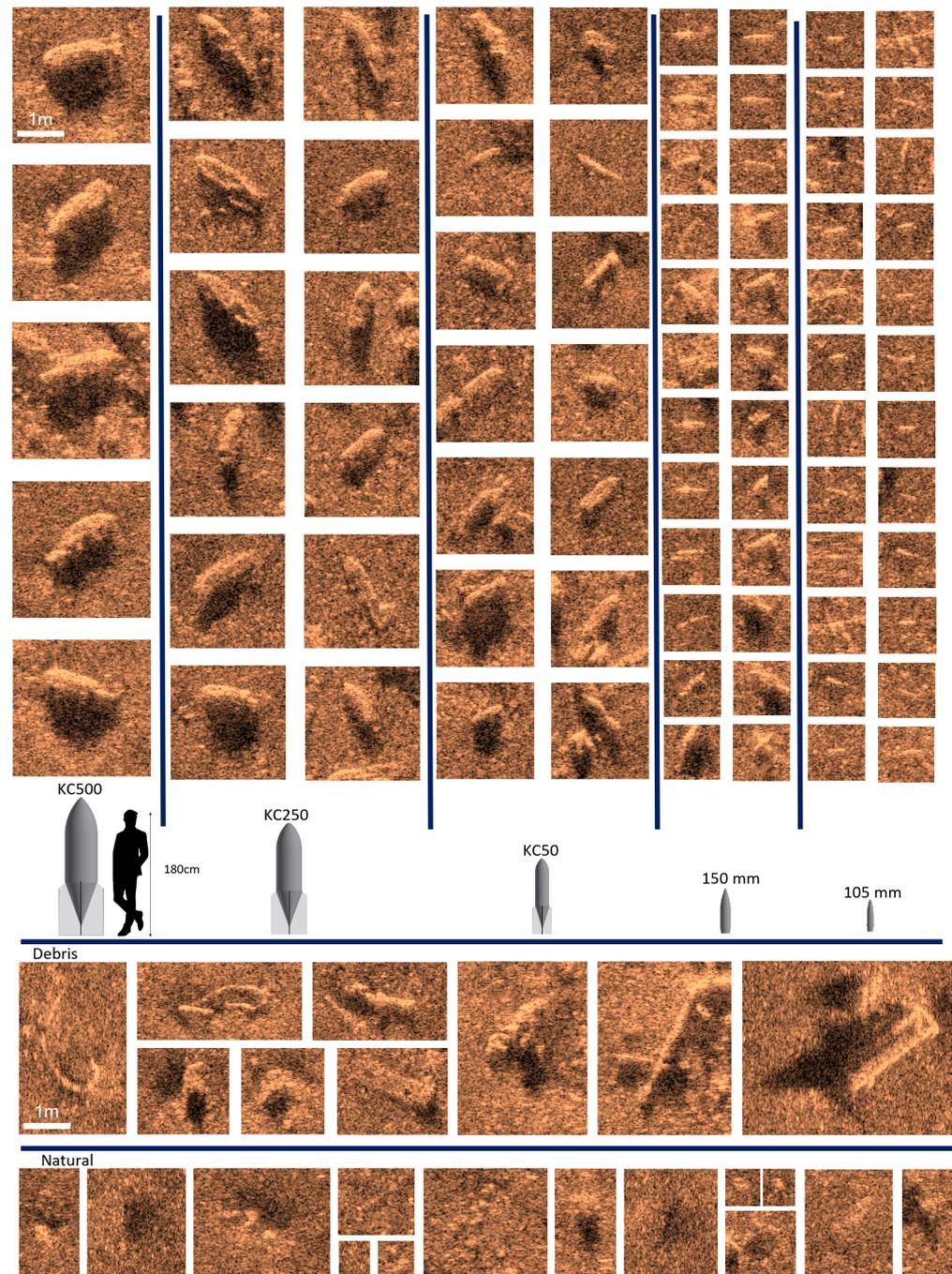


Figure 1. Objects (dangerous and harmless) found at Skagerrak, with examples from synthetic aperture sonar (all SAS images are at the same scale). Historic photos of UXOs before dumping at Skagerrak can be found in [1].

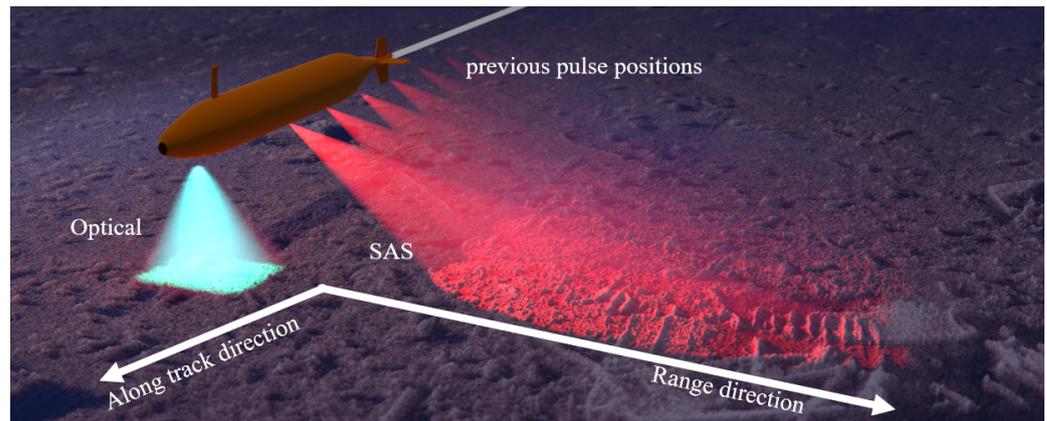


Figure 2. The geometry of data collection from an autonomous underwater vehicle (AUV). Synthetic aperture sonar (SAS) images the seafloor at a low grazing angle. Optical images are taken from a top-down orientation illuminated by a colocated light on the AUV. A typical AUV fly height for camera runs is six meters and 25–30 m for SAS runs in deep waters.

3. Labelling Methodology

The lack of ground truth in the survey area creates a need for a labelling method that incorporates uncertainty. To this end, we have used an ordinal labelling scheme, whereby each detected object is assigned one or more classes arranged in order of likelihood by the human labeller. A ranked list is a more natural and consistent process for a human labeller compared to assigning numeric values (e.g., probabilities) [13,14]. Various schemes exist for capturing object location (e.g., centre of object, bounding box, oriented bounding box, pixel mask) [15]. However, we chose to use rectangular boxes as a trade-off between precision and labelling efficiency. Note that there are automatic approaches that refine the imprecise boundaries, e.g., GrabCut [16], making labelling pixel-precise boundaries hard to justify. The entire Skagerrak survey area is 450 km², but only 0.1 km² of it was labelled, taking 35 h (more details are provided in Section 4). A single person (the lead author) performed the labelling of the multi-sensor data, via a custom-made user interface.

The labelling process was initiated on a SAS image. The image was methodically scanned, and all observed objects were bounded by a rectangular box. Each of these objects was assigned a ranked list of likely classes. The available classes are as follows and were selected based on prior knowledge of the expected chemical munition types at the dumpsite [1]:

- KC500 (500 kg bomb).
- KC250 (250 kg bomb).
- KC50 (50 kg bomb).
- 150 mm shell.
- 105 mm shell.
- Debris (anthropogenic).
- Natural.
- Not observed (available for relabelling only).

The initial labelled SAS image was used to define a common coordinate system in which all subsequent images were referenced. The bounding boxes from this previously labelled image are then overlaid onto overlapping images for independent relabelling in the different views, e.g., the optical image shown in Figure 3b. A bounding box from an alternative view or modality may not contain an observable object. Therefore, a “not observed” label was required for these cases. Any additional objects detected would also be bounded and a second pass performed on previously labelled images to label any missed objects.

Examples of two detected and labelled objects from a SAS image are given in Figure 3a. In one case, the object has been labelled as a likely 500 kg bomb, but with uncertainty due

to its resemblance to a bolder (natural). In the other case, the object has been labelled with less certainty, due to its size and shape being similar to a 50 kg bomb, a 150 mm shell, and possibly debris due to its high aspect ratio. In this case, greater detail from an overlapping optical image (Figure 3b) allowed the larger object to be labelled as a 500 kg bomb, with no feasible alternatives: high certainty. The other object, however, was not visible in the optical image so is labelled “not observed”.

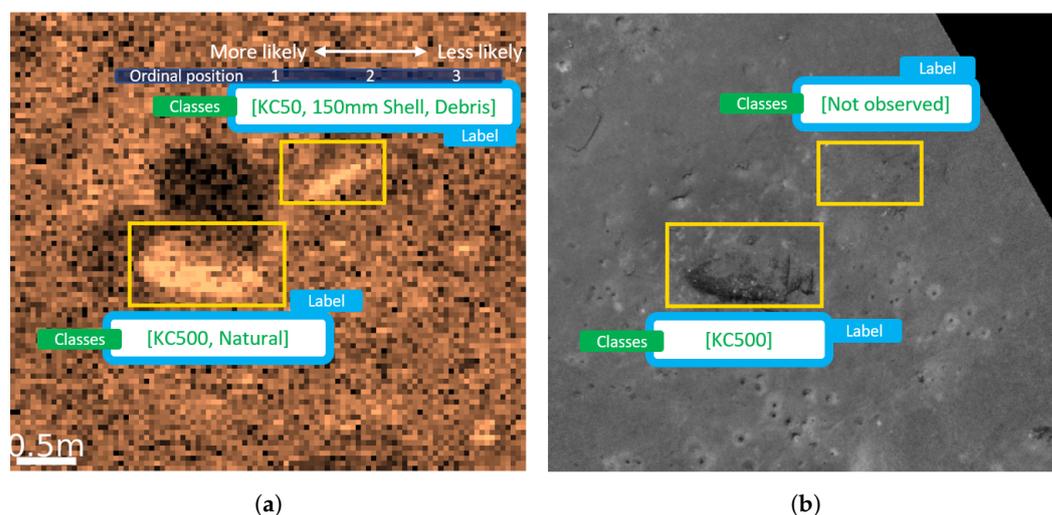


Figure 3. Object labels consist of plausible classes in order of likelihood. Class likelihood is a function of position within the label and label length. (a) SAS image with labels, and (b) optical image of the same patch of seafloor with superimposed bounding boxes relabelled. Motivation for the classes chosen is given in Section 3.

While the relabelling from different views was performed independently, the bounding box size may influence the human labeller because the size and aspect ratio implies specific UXO types. In the case of piled objects, the distinction between individual objects is not always clear, and segmentation of amorphous regions may be more suitable than the counting of individual UXOs in these areas. In summary, the probability of an object belonging to a class is some unknown function of the class’s rank (position in the ordinal list) and the length of the list. Rank reciprocal (inverse) [17], rank sum (linear) [17], and rank ordered centroid [18] are some common mappings used in multicriteria decision-making. These are consistent with our ability to rank classes. However, without a method for validating the choice of mapping the simplest regime is recommended. The multi-class, multi-view modality labels do, however, provide a rich source of information that is representative of the precision of a human labeller.

4. Results

The labelled dataset covers an area of 0.1 km²; this is only 0.02% of the total 450 km² surveyed area shown in Figure 4. However, we chose the area as it is representative of many challenging scenarios: piled objects, a variety of debris types, and sparse and dense regions. The 0.1 km² labelled area resulted in 8920 labels of 5331 objects. A single user was able to average one label every 14.2 s, taking 35 h total. Taking a “worst-case” area for labelling, as we have done, is a compromise for generating representative datasets with limited labelling resources.

The labels themselves consist of multiple classes. In Figure 5, we calculate the ratio of labels containing different combinations of classes in addition to the distribution of classes. This analysis revealed that the labeller frequently could not differentiate between objects of similar size. The labeller also found it challenging to distinguish debris, a result of debris having no expected size or shape. Historic UXOs are exposed to a corrosive environment for potentially many decades. This results in decay, biofouling, and partial or full burial, further impeding human labelling. Grouping debris and natural objects would alleviate a

frequently confused pair; however, debris is useful to classify as it indicates the presence of UXOs nearby. Grouping classes of similar objects, however, is a sensible conclusion from this result, which we investigate later.

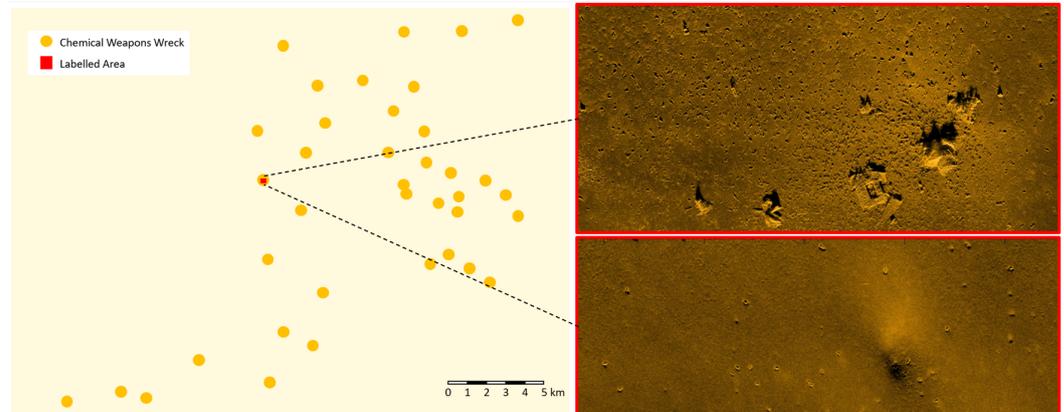


Figure 4. Suspected chemical weapons wrecks at Skagerrak [9]. The extent of the labelled area is barely visible at the scale of the full survey area. Two sonar views used for labelling are shown.

Munition types are manufactured and disposed of in varying quantities, therefore, a labelled dataset reproduces this imbalance. Figure 5b quantifies the imbalance in the Skagerrak UXO dataset with many more cases of small munitions (shells) than larger munitions (bombs). However, the number of dangerous and harmless objects is more balanced. The Imbalance between classes is a necessary consideration for the application of machine learning to avoid bias towards overrepresented classes. Practical solutions involve balancing the distribution (data-level), modifying the learning algorithm to alleviate bias (algorithm-level), or a combination of both (hybrid method) [19].

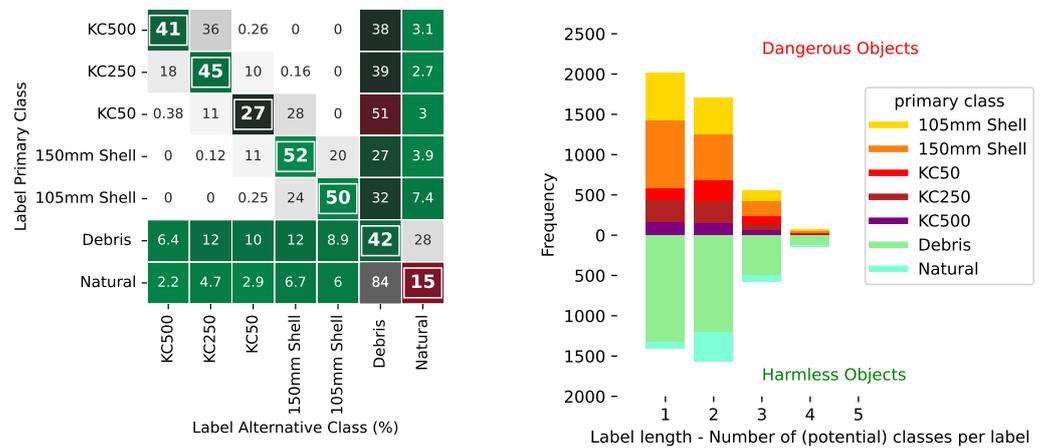


Figure 5. (a) Confusion matrix: Percentage of labels with primary class (Y-axis) that also contain (alternative) class (X-axis). Along the diagonal is the percentage of labels that contain only one class. (b) Distribution of unexploded ordnance (UXO) classes across different label lengths (uncertainty estimate).

At 100 kHz, the HISAS is capable of some penetration of sediment [2], allowing shallowly buried objects to be detected. We presupposed that the “not observed” labels for optical images correspond to buried objects. We quantify the extent of suspected burial in Figure 6 alongside examples of SAS images with their optical image counterparts taken from Skagerrak. We observed that small dense objects (shells) are more frequently buried compared to larger objects (bombs). The high proportion of buried objects is

surprising, as optical images are commonly considered as providing a means of validating SAS classifications [9].

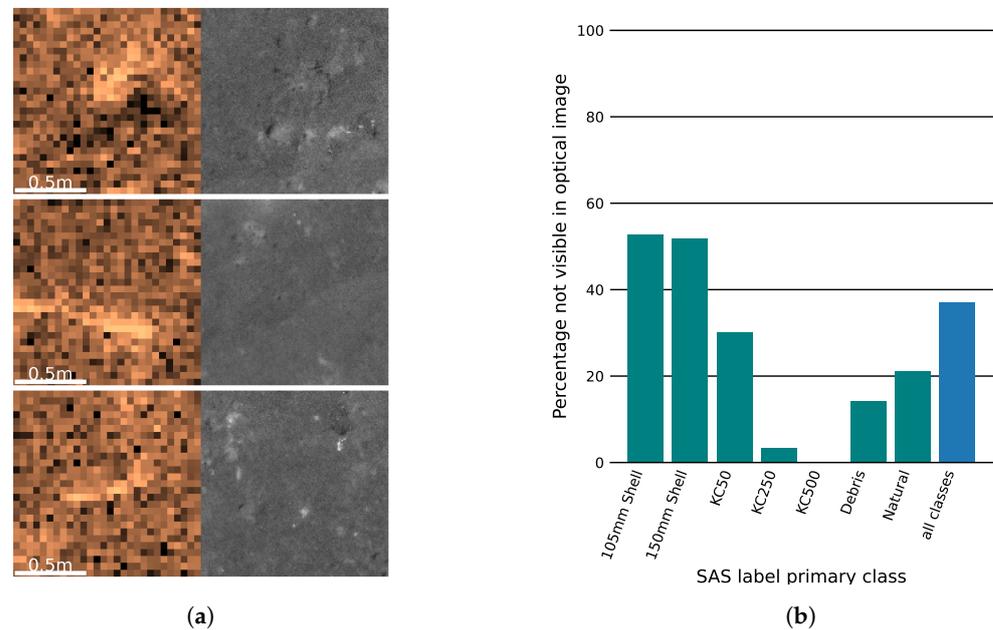


Figure 6. Observability of different objects across imaging modes: (a) The same section of seafloor imaged by SAS (visible objects) and optical (presumed buried objects). (b) Percentage of objects visible in SAS image but not optical.

The comparison between image modalities can also be made between multiple SAS views of the same objects. In Figure 7, we quantify the impact of the SAS look direction on object visibility, alongside some examples from the Skagerrak dataset. From Figure 7a, we observed that the visibility of an object's highlight and shadow can depend on its orientation. In certain cases, an object can be near-invisible from an alternative SAS look direction, and our labels show that this is more likely for smaller objects.

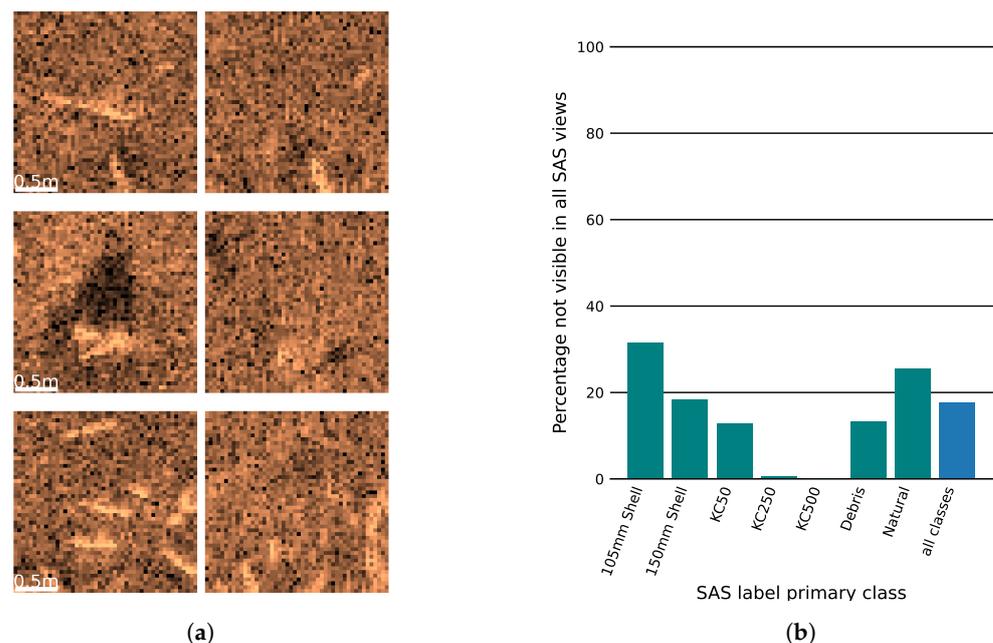


Figure 7. Observability of objects across imaging modes: (a) The same section of seafloor viewed from perpendicular SAS look directions. (b) Percentage of objects not visible in all SAS look directions (typically perpendicular).

One challenge we have identified, but are not able to quantify, is target variability: various states of assembly and decay reduce the clarity for a labeller. Instead, we quantify label accuracy, which is a consequence of target variability combined with the other challenges discussed previously. The best optical labels were identified from having only a single class assignment. The detail visible in optical images permits high confidence in these labels, which are used as pseudo ground truth. Figure 8 shows the result of assessing SAS labels against this pseudo ground truth.

Both class position and label length correlate negatively with accuracy, and this validates our use of ordinal lists for likely classes. However, 60% accuracy for the best SAS labels (green) limits the feasibility of applying supervised learning (for a comprehensive review of methods robust to noisy labels, see [20]). Furthermore, of the 5331 objects from SAS images, only 585 have optical overlap. This is generally the case due to a much slower coverage rate of $0.043 \text{ km}^2/\text{h}$ for optical images compared with $2.88 \text{ km}^2/\text{h}$ for SAS (see Section 2 for more details). Additionally, only 290 of the 585 optical images provide “pseudo ground truth” as many are obscured by partial burial and degradation. This is an insufficient number for typical supervised learning approaches [21]. The accuracy can, however, be improved by grouping into fewer classes. Using only two very broad classes of harmless versus dangerous yields an improvement in accuracy from 60% to over 80%. Separating instead into similar-sized UXOs (bombs, shells, debris, and harmless) does not significantly reduce this improvement, and provides a good trade-off for discrimination versus accuracy.

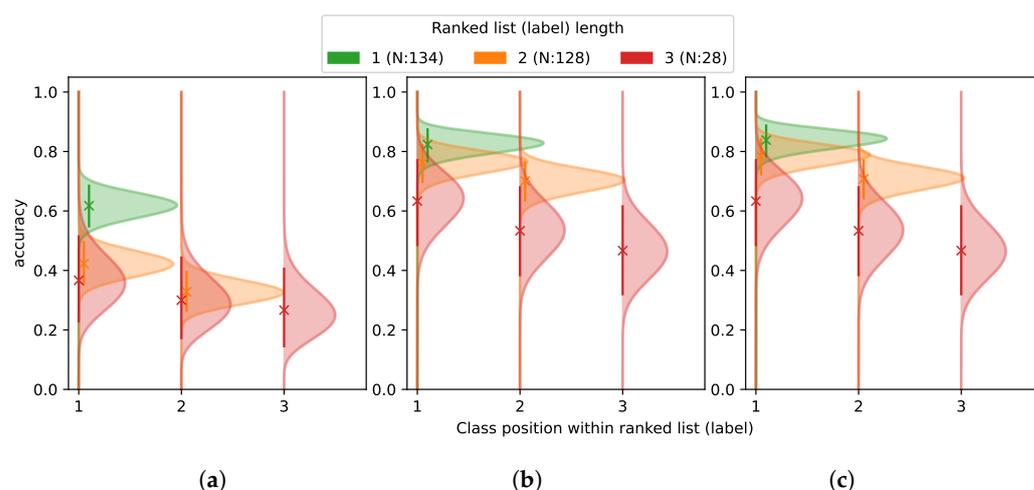


Figure 8. The Accuracy of labels compared to pseudo ground truth (best optical labels) for label length (green, orange, and red) and class position within the label (X-axis). Three different class groupings are explored: (a) individual UXO types; (b) shells, bombs, debris, and natural; (c) dangerous and harmless. Accuracy is calculated as a beta distribution from a Bernoulli likelihood (“N” successes and failures) with beta (1,1) uninformative prior.

5. Conclusions

We have defined a methodology for labelling multi-modal and multi-view images of objects on the seafloor. The methodology is intended for applications where ground truth information is unavailable, for example, at historic UXO dumpsites. It has been demonstrated on multi-view SAS and optical survey data collected from the Skagerrak chemical munitions dumpsite in Norway, where the target objects are bombs and shells amidst a field of debris and natural clutter.

A counterintuitive finding in this study was the revelation that the optical survey data do not provide sufficient ground truth of the chemical munitions dumpsite at Skagerrak for adequately labelling the objects. This is because a significant proportion of objects are buried and these are visible in the acoustic images and not the optical images. The low accuracy of labels (particularly for SAS data) and a small number of optical “pseudo

ground truth” labels motivate future research into self-supervised learning [22]. This is where a small amount of labelled data is used to bootstrap learning from a large amount of unlabelled data.

Using only optical labels with high confidence as pseudo ground truth (i.e., non-buried objects only), we quantified accuracy for the three class assignment schemes: (1) unique classes for specific UXOs (e.g., 500 kg bomb, 250 kg bomb, etc.); (2) simplified groupings of UXOs (e.g., bombs, shells, debris, and natural); and (3) a binary dangerous versus harmless label. We concluded from this work that the simplified grouping (shells, bombs, debris, and natural) gave the best trade-off between accuracy and discrimination: –60% for seven classes versus 80% for four classes.

Author Contributions: Conceptualization, O.B., R.E.H., T.S.F.H., N.W. and A.H.; data curation, R.E.H., N.W. and O.B.; methodology, O.B., R.E.H., T.S.F.H., N.W. and A.H.; formal analysis, O.B.; supervision, A.H.; writing—original draft, O.B.; writing—review and editing, O.B., R.E.H., T.S.F.H., N.W. and A.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by UK Research and Innovation (UKRI EP/S023437/1).

Data Availability Statement: The data used in this study belong to the Norwegian Defence and Research Establishment (FFI) and the Norwegian Coastal Administration. They were made available to the University of Bath with the permission of FFI. However, they cannot be made publicly available.

Acknowledgments: The authors thank the Norwegian Coastal Administration for funding the 2015 and 2016 Skagerrak data collections. The authors also thank the scientists and the crew onboard FFI’s research vessel H.U. Sverdrup II, and the HUGIN AUV operators for collecting the data in the 2015, 2016, and 2019 Skagerrak missions. The authors also acknowledge Erik Makino Bakken at the Norwegian Defence Research Establishment (FFI) for generating the camera mosaics of Skagerrak used in this study.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

SAS	Synthetic aperture sonar
AUV	Autonomous underwater vehicle
UXO	Unexploded ordnance

References

1. Arison, H.L., III. *European Disposal Operations: The Sea Disposal of Chemical Weapons*; Institute for Sea-Disposed Chemical Weapons: 2014. Available online: <https://isdcw.org/publications> (accessed on 3 September 2021).
2. Ødegård, Ø.; Hansen, R.E.; Singh, H.; Maarleveld, T.J. Archaeological use of Synthetic Aperture Sonar on deepwater wreck sites in Skagerrak. *J. Archaeol. Sci.* **2018**, *89*, 1–13. [[CrossRef](#)]
3. Lurton, X. *An Introduction to Underwater Acoustics: Principles and Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2010.
4. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [[CrossRef](#)] [[PubMed](#)]
5. Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep learning for computer vision: A brief review. *Comput. Intell. Neurosci.* **2018**, *2018*, 13. [[CrossRef](#)] [[PubMed](#)]
6. Williams, D.P.; Dugelay, S. Multi-view SAS image classification using deep learning. In *IEEE Oceans*; IEEE: Piscataway, NJ, USA, 2016; pp. 1–9.
7. Janowski, L.; Kubacka, M.; Pydyn, A.; Popek, M.; Gajewski, L. From acoustics to underwater archaeology: Deep investigation of a shallow lake using high-resolution hydroacoustics—The case of Lake Lednica, Poland. *Archaeometry* **2021**, *63*, 1059–1080. [[CrossRef](#)]
8. Czub, M.; Kotwicki, L.; Lang, T.; Sanderson, H.; Klusek, Z.; Grabowski, M.; Szubska, M.; Jakacki, J.; Andrzejewski, J.; Rak, D.; et al. Deep sea habitats in the chemical warfare dumping areas of the Baltic Sea. *Sci. Total. Environ.* **2018**, *616*, 1485–1497. [[CrossRef](#)] [[PubMed](#)]

9. Hansen, R.E.; Geilhufe, M.; Bakken, E.M.; Sæbø, T.O. Comparison of synthetic aperture sonar images and optical images of UXOS from the Skagerrak chemical munitions dumpsite. In Proceedings of the 5th Underwater Acoustics Conference (UACE2019), Heraklion, Greece, 30 June–5 July 2019; pp. 429–436.
10. Bakken, E.M.; Midtgaard, Ø. Underwater Image Mosaics for AUV-Mounted Cameras. In *Global Oceans*; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6.
11. HISAS 1030. Available online: <https://www.kongsberg.com/globalassets/maritime/km-products/product-documents/high-resolution-interferometric-synthetic-aperture-sonar-hisas> (accessed on 3 September 2021).
12. Hansen, R.E.; Sæbo, T.O.; Callow, H.J.; Hagen, P.E.; Hammerstad, E. Synthetic aperture sonar processing for the HUGIN AUV. In *Europe Oceans 2005*; IEEE: Piscataway, NJ, USA, 2005; Volume 2, pp. 1090–1094.
13. Kirkwood, C.W.; Sarin, R.K. Ranking with partial information: A method and an application. *Oper. Res.* **1985**, *33*, 38–48. [[CrossRef](#)]
14. Barron, H.; Schmidt, C.P. Sensitivity analysis of additive multiattribute value models. *Oper. Res.* **1988**, *36*, 122–127. [[CrossRef](#)]
15. Mullen, J.F., Jr.; Tanner, F.R.; Sallee, P.A. Comparing the effects of annotation type on machine learning detection performance. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019.
16. Rother, C.; Kolmogorov, V.; Blake, A. “GrabCut” interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph. (TOG)* **2004**, *23*, 309–314. [[CrossRef](#)]
17. Stillwell, W.G.; Seaver, D.A.; Edwards, W. A comparison of weight approximation techniques in multiattribute utility decision making. *Organ. Behav. Hum. Perform.* **1981**, *28*, 62–77. [[CrossRef](#)]
18. Barron, F.H. Selecting a best multiattribute alternative with partial information about attribute weights. *Acta Psychol.* **1992**, *80*, 91–103. [[CrossRef](#)]
19. Krawczyk, B. Learning from imbalanced data: open challenges and future directions. *Prog. Artif. Intell.* **2016**, *5*, 221–232. [[CrossRef](#)]
20. Song, H.; Kim, M.; Park, D.; Shin, Y.; Lee, J.G. Learning from noisy labels with deep neural networks: A survey. *arXiv* **2020**, arXiv:2007.08199.
21. Williams, D.P. Transfer learning with SAS-image convolutional neural networks for improved underwater target classification. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Yokohama, Japan, 28 July–2 August 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 78–81.
22. Jaiswal, A.; Babu, A.R.; Zadeh, M.Z.; Banerjee, D.; Makedon, F. A survey on contrastive self-supervised learning. *Technologies* **2020**, *9*, 2. [[CrossRef](#)]