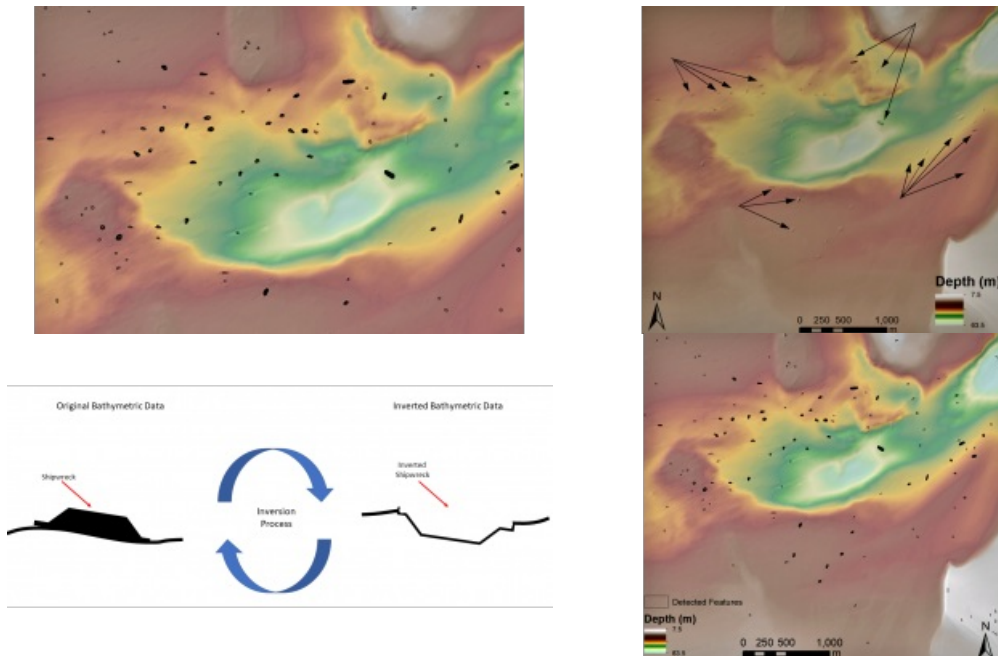*HYDROLOGICAL ALGORITHMS HELP TO UNCOVER SUNKEN HISTORY*

Automatic shipwreck detection in bathymetry data

Archaeologists have long been interested in shipwrecks. These sites can tell us about ancient transportation and trading routes, technological innovations and cultural exchanges over thousands of years. Documenting shipwrecks can be a difficult task, however. With breakthroughs in remote sensing technology (specifically sonar and radar), researchers have been able to acquire highly resolved maps of ocean floors. Consequently, we can also locate cultural objects – like shipwrecks – sitting on the bottom of oceans, lakes and other bodies of water.

Machine learning, AI and other computerized methods of analysing imagery have made incredible strides possible in many research areas. Algorithms have been devised to locate everything from buildings, roads and people to more specific objects like aeroplanes, ships and animals. Archaeologists, too, have begun using AI to locate ancient settlements. Recent research shows how the use of such automated methods can help reveal a seascape of human history sunk at the bottom of the ocean.

Bathymetry and shipwreck archaeology

Researchers have used bathymetric datasets to locate everything from ancient Roman and Greek vessels to World War II aircraft carriers. Technical studies of specific bathymetric data types have shown that shipwrecks can be identified in high-resolution imagery, but smaller or highly damaged wrecks are not as easy to locate. While most shipwreck archaeology has used bathymetric datasets produced by sonar and radar instruments, which are collected from boats or submerged vessels, some researchers have utilized bathymetric datasets collected from the air, specifically Lidar.

Lidar (light detection and ranging) data, collected using pulses of light to record 3D surfaces, has made headlines within archaeology for its ability to detect cities hidden under jungle and forest canopies. However, there are actually two kinds of Lidar data: topographic Lidar (which is used to locate features on land), and bathymetric Lidar (which can penetrate water and map shallow ocean floors). Archaeologists have used bathymetric Lidar to detect shipwrecks in shallow waters around the world. One limitation of Lidar, however, is that the water must be calm and clear for proper data collection.

The 'big data' problem and underwater archaeology

The use of 3D data like bathymetry and topography has improved researchers' ability to detect objects in image datasets. Shipwrecks, which appear as mounds at the ocean's bottom, can therefore be identified based on their depth compared to the

surrounding ocean floor. However, the ocean is vast, and recording shipwrecks and other features by hand is extremely time consuming and expensive. To this end, there are many different methods that can be used to automatically extract information from images.

The use of AI and machine learning methods have become increasingly popular as they can be extremely precise, but these methods are also very complicated and require high levels of processing power to implement. There are also other methods of pattern recognition that use a variety of object characteristics to create expectations for a computer to locate new objects. These can include colour, shape, texture, size and height, to name just a few.

Generally speaking, the greater the number of characteristics considered in a detection algorithm or model, the more accurate the detection becomes. This is particularly true if the features of interest are complex and highly variable. Different methods have had varying levels of success in detecting a range of feature types. Within archaeology, for example, automated techniques have successfully located urban complexes in Central America, Native American constructions in the United States, and tens of thousands of archaeological burial mounds and tombs across Europe and parts of Asia. Among marine archaeologists, however, these methods have not yet been extensively adopted.

Figure 1: Shipwrecks visible in bathymetric data from the National Oceanic and Atmospheric Administration in the United States. Arrows highlight some of the larger visible shipwrecks. Credit: Dylan Davis, Danielle Buffa & Amy Wrobleski (2020). (Source: https://doi.org/10.3390/heritage3020022)

For automated methods like machine learning and AI to work, image data must be at a fine enough resolution to detect patterns related to the target of interest. For example, to locate a sunken cruise liner, your data resolution can be lower than if you are looking for a canoe or kayak. For a cruise liner, your data resolution can be lower because the object is large and easy to spot, even if the imagery is not highly detailed. In contrast, something small, like a canoe or kayak, needs finely grained data to keep it from blending into the noise of coarser-grained datasets. In several countries around the world, including the United States and the United Kingdom, high-resolution bathymetric data (like sonar, radar and Lidar) is widely available. This data can allow researchers to develop automated methods for locating shipwrecks, many of which can have historical or archaeological significance (Figure 1).

Sinkhole algorithms and shipwreck detection

Because intact shipwrecks show up as mounds or lumps on the ocean floor, one particularly simple way to try and automate their detection is to use sinkhole extraction algorithms. These methods were originally developed by hydrographers to locate watersheds, sinkholes and other concave topographic features. In some of our recent research, we used a sinkhole extraction algorithm developed by geographers to detect shipwrecks located in up to 50m of water. To use these methods to locate 'mounded' features, you just need to invert the bathymetric data to turn 'mounds' into 'sinks' (Figure 2). In this way, shipwrecks become sinkholes that the algorithm can identify (Figure 3).

Figure 2: Illustration of how inverting bathymetric data works for detecting shipwrecks. The original dataset is flipped so that a mounded wreck becomes a hollow depression. The algorithm can then â€˜fillâ€™ that depression to detect a shipwreck as if it was a sinkhole or topographic depression. Credit: Dylan Davis.

Using this inversed sinkhole method, shipwrecks can be systematically identified in vast bathymetric datasets. Because the method is reproducible, researchers can use it on any dataset from any location and different researchers will produce identical results. This also helps to make research more reliable by limiting potential errors introduced by individual analysts. Especially considering the vastness of submerged areas around the world, a time-effective, reproducible method is important for underwater investigations. This is especially relevant for underwater archaeological work because submerged cultural heritage must be recorded before it is destroyed by natural and human events.

Figure 3: Example of detected shipwreck locations in the United States using sinkhole extraction algorithms. Credit: Dylan Davis, Danielle Buffa & Amy Wrobleski (2020). Source: https://doi.org/10.3390/heritage3020022.

How useful are automated methods?

Because of their complexity and often imperfect results, some may wonder if using these methods is worth it. After all, if you can scan through data and pick out shipwrecks perfectly in a couple of hours, why take all that time to train a model? Without a doubt, for researchers focusing on a limited study area, automation may not represent great use of their limited time. However, if the scale of a research project is larger – perhaps ocean-wide, or regional covering thousands of miles of coastline and open water – then the speed at which automated methods can conduct a preliminary evaluation of the data makes them worth the effort.

In an automated study of shipwreck detection conducted in the United States, we found that a sinkhole extraction algorithm could detect approximately 75% of known shipwrecks within the study region. Certainly imperfect, but a pretty good initial attempt. What was more impressive was that the method could locate shipwrecks exponentially faster than human analysis (requiring only about 3 minutes to detect over 200 shipwreck sites). When doing this by hand, it took us nearly 10 hours to record that many shipwrecks. So, even though the algorithm may have missed some shipwrecks, it saved us hours of time, and this is just for a small study area of about 20km^2.

Conclusion

Recent research has demonstrated the potential for the automated detection of shipwrecks using mathematical models. This work promises to rapidly expand the capacity of researchers to investigate the ocean floor for remnants of human history. The archaeological record is fragile and, once lost, this information can never be recovered. For submerged capsules of human history, improving our ability to detect and monitor these sites is an important undertaking, and one which computer automation is well suited to assist in.

For those interested in using automated methods for shipwreck (or other object) detection, there are many tools available within GIS systems (including open source platforms like QGIS, WhiteBox GAT and SAGA, among others), as well as some online tutorials using ArcGIS Pro and deep learning.

Further reading

1. Davis, D. S., Buffa, D. C., & Wrobleski, A. C. (2020). Assessing the utility of open-access bathymetric data for shipwreck detection in the United States. *Heritage, 3*(2), 364-383. https://doi.org/10.3390/heritage3020022
2. Zhu, B., Wang, X., Chu, Z., Yang, Y., & Shi, J. (2019). Active learning for recognition of shipwreck target in side-scan sonar image. *Remote Sensing, 11*(3), 243. https://doi.org/10.3390/rs11030243
3. Janowski, L., Kubacka, M., Pydyn, A., Popek, M., & Gajewski, L. (2021). From acoustics to underwater archaeology: Deep investigation of a shallow lake using high‐resolution hydroacoustics—The case of Lake Lednica, Poland. *Archaeometry*, arcm.12663. https://doi.org/10.1111/arcm.12663