Marine Debris: Can you beat the experts?

A sustainability collaboration between HM and TUM

Simon Chervenak, Oscar Röth, Mohamad Alkam, Zeynep Duran, and Wasuwadee Kongdech
with the help of Ekaterina Gikalo, Paul Springer, and Alexander Thamm

July 9th, 2023

In recent times, the issue of oceanic garbage has escalated into a significant and complex problem. Marine litter not only tarnishes the beauty of the sea but also poses severe threats to various aspects of our world. The detrimental impacts on the ecosystem, economy, tourism, environment, health, and the very survival of marine life cannot be ignored. This article explores the consequences of marine litter and highlights the efforts made in developing advanced algorithms that utilize satellite imagery to detect and manage waste effectively.

The consequences of oceanic garbage extend far beyond its visual impact. The fragile balance of marine ecosystems suffers tremendously as a result. Aquatic organisms mistakenly consume or become entangled in the debris, leading to injury, suffocation, or death. Additionally, the chemicals and toxins present in the litter can contaminate the water, further disrupting the delicate web of life beneath the waves. Furthermore, the economic and tourism sectors also experience adverse effects. Coastal areas that once thrived on the allure of pristine beaches now face the challenge of dealing with immense amounts of waste washing ashore. This has a direct impact on tourism, discouraging visitors and resulting in financial losses for local businesses. Additionally, the cleanup efforts required to maintain cleanliness along coastlines place a burden on already strained economies.

The environment at large suffers as well, as the accumulation of garbage in the oceans contributes to climate change. Plastics, for instance, break down into microplastics, which enter the food chain and eventually find their way onto our dinner plates. This vicious cycle poses a threat to human health and raises concerns about the long-term consequences of our actions.

Eliminating the threat of marine litter to our ecosystem is a complicated and multi-layered process. However, at the moment we rely on humans to find and remove marine litter without a global detection system in place. Artificial intelligence can bridge the gap and be a linchpin in this system. By analyzing satellite images, a developed artificial intelligence system could detect the litter and forward this information to boats across the globe. Armed with this knowledge, the oceans could be cleaned once and for all.

An easy task for an AI to start with would be detecting floating litter, since this litter can be picked up on satellite cameras. In order to train such models, and contribute to worldwide data collection, scientists created the MARIDA dataset, which takes pictures like Fig. 1 and meticulously annotates each pixel to tell the AI if it is litter or water. From this, the AI can learn and develop a detection algorithm for detecting floating marine litter. These images come from twelve locations across the globe to attempt to teach the AI how to recognize litter in diverse locations. The researchers also included classes such as sargassum and turbid water that are ignored for the purposes of this experiment.

Fig 1. Marine litter floating on the Pacific. (Aguilera)

Technological Innovations: Detecting Marine Debris from Satellite Images

Realizing the gravity of the situation, scientists and researchers have been diligently working to find solutions to combat marine litter. Among the most promising innovations is the use of advanced algorithms that leverage satellite imagery to detect and manage waste more efficiently.

Through extensive research and development, these algorithms have been designed to analyze satellite images and identify patterns indicative of marine debris. By automating this process, large-scale detection and monitoring of oceanic garbage become feasible. The algorithms can accurately distinguish between natural features and floating debris, providing valuable data for waste management initiatives. Implementing such technological solutions offers several advantages. First and foremost, it enables a comprehensive understanding of the extent and distribution of marine litter, allowing for targeted cleanup operations. By pinpointing the areas most affected by garbage accumulation, resources can be allocated effectively, maximizing the impact of cleanup efforts. Moreover, these algorithms assist in assessing the effectiveness of waste management strategies. By tracking changes in debris patterns over time, policymakers and researchers can gauge the success of implemented measures and adapt accordingly. The availability of accurate and up-to-date data aids in evidence-based decision-making, facilitating proactive interventions to address the issue at its core.

Furthermore, technological advancements in detecting marine debris from satellite images can play a crucial role in raising public awareness. The visual representation of the extent of oceanic garbage captured by satellite imagery can be a powerful tool to educate and engage individuals, communities, and policymakers. By fostering a sense of urgency and responsibility, we can mobilize collective efforts to reduce, prevent, and ultimately eliminate marine litter.

Improving the Model

When it came to improving the model, we had a few main ideas. The researchers proposed several models; we chose to improve the most promising one, the UNet model (Ronneberger, 2015). The UNet model was proposed for the task of biomedical image segmentation, where different pixels need to be separated into classes. It is a fast and ideal network for training on annotated pixels, as in the case of the MARIDA dataset.

Our first improvement was to look at the data that the model was running on. Data for training a neural network is split into three parts (Myrianthous, 2021). The training set is the data the network trains and learns on and comprises the majority of the data. The validation and testing sets are small parts of the data held back to test the models' accuracy on unseen data. The researchers randomly split up the data, which did not account for the spatial nature of the data. Since each image comes from a different location around the globe, and each location can look different, we decided to split up the sets to acknowledge this. We chose two of the twelve locations and put them both entirely in the training set to see if the model could generalize to these new unseen locations.

The researchers set up a basic UNet architecture with a standard Adam optimizer and cross-entropy loss function. The UNet architecture is a type of neural network designed for differentiating parts of images. It takes in an image file as input and outputs a mapping of how it divided the image into the different classes (water, litter, etc). To train a UNet, the researcher feeds it hundreds of satellite images for which the researchers know the answer (from the annotation). The AI compares the answer of its algorithm to the known answer and adjusts its parameters accordingly. After looking at all of the images, the AI is then tested on the testing set to determine its accuracy.

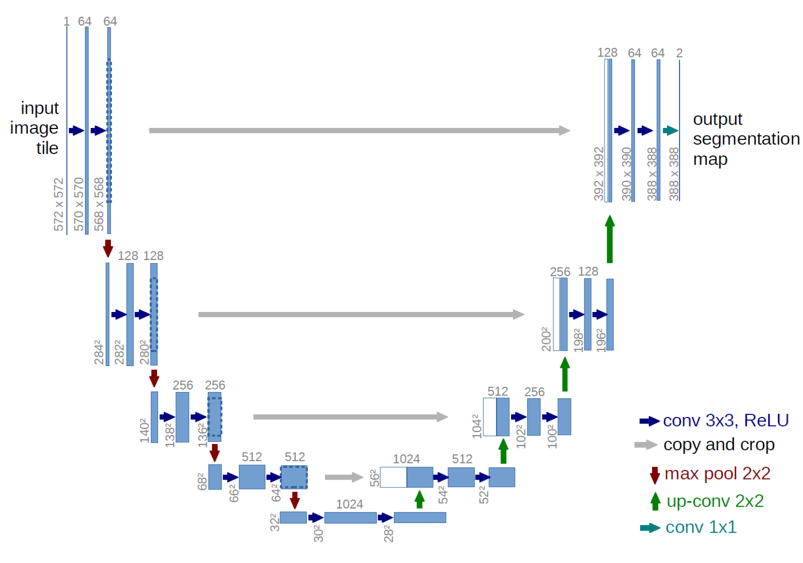


Fig 2. A basic UNet model. (Ronneberger, 2015).

The UNet resembles a U-shape, which represents the way it handles images. The images are downsampled (scaled to a smaller size) to emphasize the important parts of the image, and then upsampled to return a segmentation map of the size of the original image. This helps the algorithm only focus on the important parts of the image rather than get confused from looking at the entire image.

Our next optimization was to change this loss function to the newly proposed focal loss function (Lin et al., 2018). This function, specialized for semantic segmentation, helps the model prioritize classes which are underrepresented in the data set and harder, more difficult to classify examples. It is a modification of the cross-entropy loss function that the researchers were already using. This function is perfect for detecting small objects with large quantities of easily distinguished background examples, which is why we chose it to detect the tiny quantity of marine litter pixels among the background of water pixels.

Another idea for improving the model's accuracy was to reduce the number of classes. Indeed, as the aim of the model is to detect marine litter, the fact that there are 12 classes means that the model has to decide between 12 possibilities each time, thus increasing the computation time (decision) but also potentially inducing an increase in decision errors. Initially, we wanted to create a binary system (the pixel is marine litter or not) but we then chose to reduce the model to 5 classes by grouping together classes that designate the same things. For example, to aggregate the different water classes (Mixed Water to Marine Water Class, Wakes to Marine Water Class, Cloud Shadows to Marine Water Class, Waves to Marine Water Class) into the same class that distinguishes water in general. We thus reduced the number of classes to 5.

We then tested this model with the data, but this did not bring any significant improvement to the model.

To visualize the data on a map, we extracted the coordinates of each location, and then created a csv file containing these coordinates. Once these coordinates had been extracted, we sorted them by cluster. As a result, it was quite easy to draw a map with the different clusters.

This data extraction not only enabled data visualization, but also the implementation of other ideas to improve the model. Instead of having a random distribution of training, test and validation data, we opted for an equitable distribution between the different sites. This means that data from each cluster appeared proportionally to the other in the training, test and validation dataset. The idea behind this implementation was that the model is not biased by over-representation of one cluster and effectively ignores the others, ensuring a fair representation of the data in the model. But once again, improvements to the model were unsatisfactory.

Our final idea for improvement was to change the architecture of the model. We tried two different architectures. Our first attempt was to make the given UNet wider by adding another layer of nodes, but this had a negligible effect on accuracy. There is some evidence that deeper UNets can forget information that they learned in earlier layers, so this may not have been the best idea. We also attempted to harness the DeepLab ResNet architecture developed by Google (Mao, 2018) for real-time segmentation for autonomous vehicles. However, this network did not harness the non-RGB layers that the satellite data provided, and did not perform particularly well on the data.

These improvements did yield slight increases in both the precision and recall of the neural network with respect to marine debris. Further improvements can definitely occur, particularly from those with special knowledge about manipulating the eight non-visible-light channels contained within the data. Without this knowledge or further data manipulation, however, the dataset is ultimately subpar for such an undertaking, as there are simply not enough annotated marine litter pixels. Even a model that identifies 1% of the water pixels as marine litter (such as the DeepLab architecture) still results in a precipitous drop in accuracy that makes it seem that the model is essentially guessing.

Data Visualization

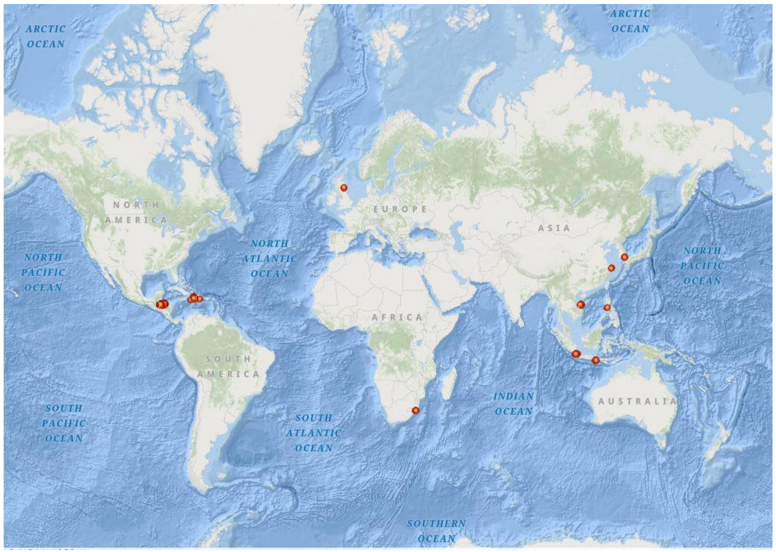
Data is critical to decision making. Trustworthy information is essential. Data visualization means integrating and combining data from various sources and making it accessible as a single data store.

By being like a single database, data visualization is a technology that allows data to be managed, integrated and analyzed. Visualization of stored data in image or video format has a very important role in this regard.

Data visualization can be done using special tools or special software. Data visualization servers can be accessed through a web-based interface. The process of data visualization consists of several steps. Firstly, data sources are defined. Secondly, connection to the data sources and also extraction of the data that needs to be visualized. After the data is extracted, the data is transformed and cleansed in order to make it usable.

The main point in data visualization is to create the visual data layer. The visual data layer provides transformation to the appropriate format for the usable data sources. Data visualization can be used to create charts, graphs, maps or dashboards (Jenifa, 2022).

In this project, ArcGIS Data and Maps software is used to access the data layers for the visualization of data. This software is a collection of data layers, including boundary, demographic, basemap, and thematic layers for the US, Europe, and all the world ('ArcGIS Data and Maps'). After the data was uploaded to the system, the data was visualized on the map by labeling the coordinates where the marine debris were located.



Conclusion

The problem of oceanic garbage poses significant threats to our ecosystems, economies, tourism, environment, health, and marine life. However, advancements in technology offer hope for effective waste management solutions. Algorithms utilizing satellite imagery for detecting marine debris provide crucial insights, enabling targeted cleanup operations, assessing the effectiveness of strategies, and raising public awareness. By harnessing the power of innovation and collaboration, we can work towards a cleaner, healthier, and more sustainable future for our oceans. Although we have made some improvements to the model's accuracy, a more thorough and holistic examination of the channels of data provided is necessary to unlock the true potential of the MARIDA dataset.

References

Aguilera, M. (2009, August 28). Growth of ocean 'garbage patch' alarms experts. NBC News. https://www.nbcnews.com/id/wbna32592123

'ArcGIS Data and Maps'. *ArcGIS*, https://www.arcgis.com/home/group.html?id=24838c2d95e14dd18c25e9bad55a7f82#overview.

Jenifa, A. (2022, December 26).'What Is Data Virtualization and Why Do We Need It'. *Geekflare*, https://geekflare.com/data-virtualization/.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2018). Focal loss for dense object detection. arXiv. https://doi.org/10.48550/arXiv.1708.02002

Mao, L. (2018, December 11). Semantic segmentation using deeplab v3. Lei Mao's Log Book. https://leimao.github.io/project/DeepLab-V3/

Myrianthous, G. (2021, August 3). Training vs testing vs validation sets. Medium. https://towardsdatascience.com/training-vs-testing-vs-validation-sets-a44bed52a0e1

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation (Version 1). arXiv. https://doi.org/10.48550/arXiv.1505.04597